

100-34  
406-46  
p- 91

# Hybrid Upwind Splitting (HUS) by a Field-by-Field Decomposition

N95-21957

Unclas

G3/34 0040646

Frédéric Coquel  
*Institute for Computational Mechanics in Propulsion*  
*Lewis Research Center*  
*Cleveland, Ohio*

and

Meng-Sing Liou  
*National Aeronautics and Space Administration*  
*Lewis Research Center*  
*Cleveland, Ohio*

January 1995

(NASA-TM-106843) HYBRID UPWIND  
SPLITTING (HUS) BY A FIELD-BY-FIELD  
DECOMPOSITION (NASA. Lewis  
Research Center) 91 p



National Aeronautics and  
Space Administration





# HYBRID UPWIND SPLITTING (HUS) BY A FIELD-BY-FIELD DECOMPOSITION

Frédéric Coquel<sup>†</sup> and Meng-Sing Liou<sup>‡</sup>

## Abstract

We introduce and develop in this paper a new approach for upwind biasing : the *Hybrid Upwind Splitting* (HUS) method. This original procedure is based on a suitable hybridization of current prominent *Flux Vector Splitting* (FVS) and *Flux Difference Splitting* (FDS) methods. The HUS method is designed to naturally combine the respective strengths of the above methods while excluding their main deficiencies. Specifically, the HUS strategy yields a family of upwind methods that exhibit the *robustness of FVS schemes in the capture of nonlinear waves* and the *accuracy of some FDS schemes in the resolution of linear waves*.

We give a detailed construction of the HUS methods following a *general and systematic procedure* directly performed at the basic level of the *Field by Field* (i.e. waves) decomposition involved in FDS methods. For such a given decomposition, each field is endowed either with FVS or FDS numerical fluxes, depending on the nonlinear nature of the field under consideration. Such a design principle is made possible thanks to the introduction of a convenient formalism that provides us with a unified framework for upwind methods.

The HUS methods we propose bring significant improvements over current methods in terms of accuracy and robustness. They yield *entropy-satisfying* approximate solutions as they are strongly supported in numerical experiments. Field by field hybrid numerical fluxes also achieve *fairly simple and explicit* expressions and hence require a computational effort between that of the FVS and FDS.

Several numerical experiments ranging from stiff 1D shock-tube to high speed viscous flows problems are displayed, intending to illustrate the benefits of the present approach. We shall assess in particular the relevance of our HUS schemes to viscous flow calculations.

---

<sup>†</sup> Institute for Computational Mechanics in Propulsion (ICOMP), NASA Lewis Research Center, Cleveland, OH 44135; Permanent address: Theoretical Aerodynamics Branch 1, ONERA, B.P 72 92322 Châtillon Cedex, France.

<sup>‡</sup> Internal Fluid Mechanics Division, NASA Lewis Research Center, Cleveland, OH 44135.

## Introduction

Over the past decades, the research effort in computational fluid dynamics (CFD) has resulted in several prominent and widely used schemes. In particular, two distinct and complementary approaches for upwind biasing have been proposed and investigated, namely the Flux Vector and Flux Difference Splitting methods (FVS and FDS respectively). Each has demonstrated fairly interesting capabilities, inherited from their design principle and assumptions, which are manifested in the numerical calculation in the form of **either robustness or accuracy**. In other words, the design principle directly determines the effectiveness of the method; abundant in the literature are papers reporting their successes and failures in calculations.

The first approach (FVS) has received attention mainly for the Euler equations. Harten et al. [22] referred it as the Boltzmann approach for its formal links with the gas dynamic setting where the flow is modeled by a set of pseudo-particles evolving freely according to a collisionless Boltzmann equation. As a consequence of the resulting mixing of particles, FVS methods suffer from a constitutive lack of resolution for *linear waves* [50] (contact discontinuities), thereby making them irrelevant candidates for Navier-Stokes calculations since the viscous layer behaves in the limit like a contact discontinuity. On the other hand, FVS methods have demonstrated their ability to capture intense and only admissible *nonlinear waves* (such as shock waves).

The second approach (FDS), generally classified as approximate Riemann solvers, has been developed primarily to provide approximate information regarding wave speeds and amplitudes in the Riemann problem. Two most outstanding FDS methods are proposed by Osher [35] and Roe [45]. These FDS methods specifically provide a direct representation of the linearly degenerate field with which the contact discontinuity is associated. Since they both, by construction, reproduce the exact solution for a single stationary contact discontinuity, they become accurate and relevant candidates for solving the Navier-Stokes (at least the steady) equations. Compared to Godunov's exact solution of the Riemann problem [17], Osher's scheme replaces the shock by an overturned compression wave and thus makes an explicit computation of numerical fluxes possible, but at the expense of reducing the robustness. Jacob [24] reported a lack of robustness in the Osher splitting for capturing a strong detached shock.

More drastic simplification is used in the Roe splitting, rendering it an even cheaper procedure in which the transonic expansion is replaced by a shock, but with the consequence of allowing a physically inadmissible expansion shock. Quirk [41] recently summarized some failings encountered by using the Roe splitting in calculating *nonlinear*

*waves.*

Therefore, FVS and FDS approaches do not yield the same advantages and drawbacks, but they clearly complement each other.

These observations have prompted some recent works in implementing some switches directly aimed at choosing either a FVS or a FDS method according to the local smoothness of the solution but at the expense of tunable parameters and of a discontinuous method (see [41] for a related idea).

Up to our knowledge, the first scheme combining both FVS and FDS features into a single flux for the sake of robustness and accuracy is the Advection Upstream Splitting Method (AUSM) introduced by Liou and Steffen [30]. This original splitting is indeed based on the appropriate definition of a cell-face advection Mach number. Liou [32] has recently proposed a convenient framework for the development of a family of schemes, referred to as AUSM+, which further improves AUSM while preserving its advantageous features. Blending flux vector- and difference-splittings in the AUSM, Wada and Liou [51] have achieved a promising scheme, termed AUSMDV, which also cures difficulties found in each constituent scheme.

In this paper, we propose another approach for upwind splitting : namely the field by field Hybrid Upwind Splitting. The formalism blends in a natural way the complementary properties exhibited by FVS and FDS methods. Indeed, the procedure we propose yields upwind methods built to share the very *robustness of FVS schemes in the capture of nonlinear waves* and the *accuracy of some FDS schemes in the resolution of linear waves*. These new schemes are derived following a general hybridization technique directly performed at the basic level of the *field by field* decomposition involved in FDS methods. In particular, we stress that our hybrid upwind splitting is free of tunable parameters.

The format of this paper is as follows. In Section 2, we introduce a convenient formalism for upwind methods that provides us with a “unified” framework for upwinding. Our framework turns out to be sufficiently large to recover the FDS methods and at the same time the FVS ones. Furthermore, its formalism enables a natural distinction between these two approaches. Roughly speaking, we shall see in what sense FVS methods completely ignore the concept of field by field decomposition which is precisely the root of all the FDS methods. This rather obvious distinction is revisited here for its deep consequences, leading to in particular a very natural guideline for the design of our Hybrid Upwind Splitting methods.

Derived in full details in Section 3, these new methods fall into the framework we have developed in Section 2. Some of their features will be analyzed there. Section 4 is especially devoted to the applications to the Euler equations. Different HUS methods of interest are given an explicit and self-contained form of expression. Finally, we present in Section 5 several numerical experiments directly devoted to measuring the relevance of the HUS methods proposed hereinafter.

## I. Notation

In this paper, we consider numerical approximations of weak solutions to the following Cauchy problem for nonlinear hyperbolic systems of conservation laws

$$\frac{\partial u}{\partial t} + \frac{\partial \mathbf{f}(u)}{\partial x} = 0, \quad t > 0, \quad x \in \mathbb{R}, \quad (1.1)$$

$$u(0, x) = u_0(x). \quad (1.2)$$

Here,  $u(t, x)$  belongs to a phase space  $\mathcal{U} \in \mathbb{R}^p$  and  $\mathbf{f}$  denotes a smooth flux function,  $\mathbf{f} : \mathcal{U} \rightarrow \mathbb{R}^p$ , whose Jacobian matrix  $A(u) = d\mathbf{f}(u)$  is assumed to be diagonalizable with  $p$  real eigenvalues  $\lambda_k(u)$ ,  $1 \leq k \leq p$ . As is well known, this problem in general does not admit smooth solutions so that weak solutions in the sense of distributions must be considered. Moreover, an entropy-satisfying condition must be added to select the physically relevant (or entropy) weak solutions. We refer to Lax [28] and to Godlewsky and Raviart [16] for precise statements and detailed results.

We shall assume in the sequel that (1.1) is strictly hyperbolic, that is the Jacobian matrix  $A(u)$  admits  $p$  distinct eigenvalues. The right eigenvectors, corresponding to eigenvalues arranged in increasing order, are denoted by

$$r_1(u), r_2(u), \dots, r_p(u), \quad (1.3)$$

while

$$l_1^T(u), l_2^T(u), \dots, l_p^T(u), \quad (1.4)$$

refer to the associated left eigenvectors.

The mapping  $u \in \mathcal{U} \rightarrow r_k(u)$  is called the  $k$ th field. Throughout this paper, we shall assume that all the fields are either genuinely nonlinear in the sense of Lax [28], i.e. (after a suitable normalization)

$$\nabla \lambda_k(u) \cdot r_k(u) = 1, \quad \forall u \in \mathcal{U}, \quad (1.5a)$$

or linearly degenerate (see [28]), i.e.

$$\nabla \lambda_k(u) \cdot r_k(u) = 0, \quad \forall u \in \mathcal{U}. \quad (1.5b)$$

The following definitions, although standard, will be of constant use in the sequel. If the  $k$ th field is genuinely nonlinear, a  $k$  shock wave moving at speed  $\sigma$  is defined to be a discontinuity of  $u$  subject to the following Rankine-Hugoniot condition

$$\sigma(u_R - u_L) = (f(u_R) - f(u_L)). \quad (1.6)$$

This  $k$  shock is said, after Lax [28], to be admissible if

$$\lambda_k(u_L) - \sigma > 0 > \lambda_k(u_R) - \sigma. \quad (1.7)$$

Associated with the  $k$ th genuinely nonlinear field, a  $k$  rarefaction wave is defined to be a solution of the following ordinary differential equation in the phase space  $\mathcal{U}$

$$\frac{du}{ds} = r_k(u(s)), \quad s \in \mathcal{R}. \quad (1.8)$$

In view of requirement (1.5a), note that the parameter  $s$  in (1.8) is in fact restricted to an open subset of  $\mathcal{R}$  where the mapping  $s \rightarrow \lambda_k(u(s))$  strictly increases along the integral curve of (1.8) oriented by  $r_k$ .

As for the linearly degenerate field, a  $k$  contact discontinuity is defined to be a discontinuity that satisfies (1.6) as well but with  $\sigma = \lambda_k(u_L) = \lambda_k(u_R)$ , or a  $k$ th wave along which the differential equation (1.8) is trivially satisfied.

In the following, we shall describe numerical approximations to weak solutions of (1.1). Let  $h$  and  $h_x$  be respectively the time and space steps. We consider piecewise constant approximate solutions  $u^h : R_+ \times R \rightarrow \mathcal{U} \in \mathcal{R}^p$ ; i.e., for  $t \in [nh, (n+1)h]$ ,  $n \in N$ ,

$$u^h(t, x) = u_i^n, \quad x \in [(i - \frac{1}{2})h_x, (i + \frac{1}{2})h_x[, \quad i \in Z, \quad (1.9)$$

with

$$u_i^0 \equiv \frac{1}{h_x} \int_{(i-\frac{1}{2})h_x}^{(i+\frac{1}{2})h_x} u_0(x) dx, \quad i \in Z. \quad (1.10)$$

For  $n > 1$ ,  $(u_i^{n+1})_{i \in Z}$  is defined by the following 3-point explicit scheme in conservation form

$$u_i^{n+1} = u_i^n - \frac{h}{h_x} (f_{i+\frac{1}{2}}^n - f_{i-\frac{1}{2}}^n), \quad i \in Z. \quad (1.11)$$

Here, the numerical flux defined by

$$f_{i+\frac{1}{2}} = f(u_i^n, u_{i+1}^n) \quad (1.12)$$

is assumed to satisfy the following smoothness and consistency conditions :

$$f : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}^p \text{ is a locally Lipschitz - continuous function} \quad (1.13)$$

$$f(u, u) = \mathbf{f}(u), \quad \forall u \in \mathcal{U}. \quad (1.14)$$

By the well known theorem of Lax and Wendroff (see [20]), if the limits of approximate solutions (1.9), as  $h \rightarrow 0$ , exist a.e. in the sense of bounded  $L^1_{loc}$  convergence and provided that the assumptions (1.13)-(1.14) are satisfied, then these limits are weak solutions of (1.1). We refer to [16] for an exhaustive presentation and the updated references therein.

## II. A General Setting for Upwind Splitting Methods

In this section, we introduce a convenient formalism for upwind methods. This formalism enables the classical FDS and FVS methods to be re-expressed in the same framework and furthermore it encompasses a new and general concept for upwind biasing (see Section III).

In the sequel, we shall focus our attention on upwind schemes in the sense of the framework we now introduce. This framework is based on the fruitful notion of paths introduced by Dal Maso, Le Floch and Murat [9] to deal with nonconservative products arising in a hyperbolic system in nonconservation form. Following their definition, we consider a fixed family of paths  $\Phi$  in  $\mathcal{U}$ , that is a map  $\Phi : [0, 1] \times \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{U} \in \mathcal{R}^p$  assumed to satisfy the following properties of consistency and smoothness,

$$\Phi(0; u_L, u_R) = u_L \text{ and } \Phi(1; u_L, u_R) = u_R, \quad \text{for any } u_L \text{ and } u_R \text{ in } \mathcal{U}, \quad (2.1a)$$

$$\Phi(s; u, u) = u, \quad \text{for any } u \text{ in } \mathcal{U} \text{ and any } s \text{ in } [0, 1], \quad (2.1b)$$

and

*For every bounded set  $K$  in  $\mathcal{U}$ , there exists a constant  $C > 0$  such that for any states  $(u_L, u_R)$  and  $(v_L, v_R)$  in  $K$  :*

$$\left| \frac{\partial \Phi(s; u_L, u_R)}{\partial s} - \frac{\partial \Phi(s; v_L, v_R)}{\partial s} \right| \leq C |(u_R - u_L) - (v_R - v_L)|, \quad \text{a.e. } s \text{ in } [0, 1]. \quad (2.1c)$$

As an immediate consequence of (2.1c), we have for every  $u_L$  and  $u_R$  in  $K$

$$\left| \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right| \leq C |u_R - u_L|, \quad \text{a.e. } s \text{ in } [0, 1]. \quad (2.1d)$$



We shall need also in the sequel to consider matrix-valued functions  $B^+$  and  $B^- : \mathcal{U} \times [0, 1] \rightarrow \text{Mat}(\mathcal{R}^p, \mathcal{R}^p)$  such that :

$$\begin{aligned} &B^+(u, s) \text{ (respectively } B^-(u, s)) \text{ is diagonalizable with only} \\ &\text{real and nonnegative } \lambda^+(u, s) \text{ (resp. real and nonpositive } \lambda^-(u, s)) \\ &\text{eigenvalues for all } u \text{ in } \mathcal{U} \text{ and any given } s \text{ in } [0, 1]; \end{aligned} \quad (2.2a)$$

and

$$\begin{aligned} &\text{For every bounded set } K \text{ in } \mathcal{U}, \text{ there exists a constant } C > 0 \text{ such that,} \\ &\text{for all } u \text{ and } v \text{ in } K, \text{ a.e. } s \text{ in } [0, 1], \\ &\|B^+(u, s)\| \leq C, \quad \|B^-(u, s)\| \leq C; \end{aligned} \quad (2.2b)$$

$$\|B^+(u, s) - B^+(v, s)\| \leq C|u - v|, \quad \|B^-(u, s) - B^-(v, s)\| \leq C|u - v|. \quad (2.2c)$$

Based on a family of paths  $\Phi$  and two matrix-valued functions  $B^\pm$ , we propose the following framework for upwind biasing.

### Definition and Proposition 1

Let  $\Phi$  be a fixed family of paths in the sense of (2.1) and let  $B^+$  and  $B^-$  be two matrix-valued functions satisfying (2.2).  $\Phi$  and  $B^\pm$  are assumed to satisfy the following compatibility condition

$$\int_0^1 \left( B^+(\Phi(s; u_L, u_R), s) + B^-(\Phi(s; u_L, u_R), s) \right) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = \mathbf{f}(u_R) - \mathbf{f}(u_L). \quad (2.3)$$

Under the CFL-like condition

$$\frac{h}{h_x} \max_{1 \leq k \leq p} |\lambda_k^\pm(u, s)| \leq \frac{1}{2}, \quad (2.4)$$

let us now consider the following two averaging

$$\tilde{u}_L(u_L, u_R) = u_L - \frac{2h}{h_x} \int_0^1 B^-(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds, \quad (2.5a)$$

$$\tilde{u}_R(u_L, u_R) = u_R - \frac{2h}{h_x} \int_0^1 B^+(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds. \quad (2.5b)$$

We define a  $(\Phi; B^\pm)$  scheme as follows :

$$u_i^{n+1} = \frac{1}{2} \left( \tilde{u}_R(u_{i-1}^n, u_i^n) + \tilde{u}_L(u_i^n, u_{i+1}^n) \right). \quad (2.6)$$

Assertion. This definition yields a consistent 3-point scheme in conservation form for which the associated numerical flux is written :

$$f(u_L, u_R) = \frac{1}{2} \left( f(u_L) + f(u_R) - \int_0^1 \left\{ B^+(\Phi(s; u_L, u_R), s) - B^-(\Phi(s; u_L, u_R), s) \right\} \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds \right). \quad (2.7)$$

### Remarks

2.1- The averaging (2.5a) may be understood as the integral conservation law applied over the rectangle  $(-h_x/2, 0) \times (0, h)$  to the approximate solution of a specified Riemann problem (see the CFL-like condition (2.4)). Conversely, the averaging (2.5b) has the same physical picture but using the rectangle  $(0, h_x/2) \times (0, h)$ . As a consequence, formula (2.6) can be seen as the averaging under the CFL condition (2.4) of two neighboring and non-interacting half approximate-Riemann solutions. Such a formal interpretation obviously applies to the FDS methods (we refer the reader to Remark 2.6 concerning the FVS schemes).

2.2- Requirement (2.3) is an essential compatibility condition to be met by the family of paths  $\Phi$  and the  $B^\pm$  functions with respect to the problem (1.1). Indeed, summing (2.5a) and (2.5b) gives :

$$\frac{1}{2}(\tilde{u}_L + \tilde{u}_R) = \frac{1}{2}(u_L + u_R) - \frac{h}{h_x} \int_0^1 \{B^+ + B^-\}(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds. \quad (2.8a)$$

This relation can be thought of as applying the integral conservation law over the rectangle  $(-h_x/2, h_x/2) \times (0, h)$  to the approximate solution quoted in the above Remark 2.1. Under the CFL condition (2.4), the compatibility requirement (2.3) clearly enforces the above averaging to coincide with the relevant one provided by the Godunov exact Riemann solver :

$$\frac{1}{2}(\tilde{u}_L + \tilde{u}_R) = \frac{1}{2}(u_L + u_R) - \frac{h}{h_x} (f(u_R) - f(u_L)). \quad (2.8b)$$

In view of (2.8), the compatibility condition (2.3) might be also read as a consistency requirement with respect to (1.1). In fact, (2.3) yields a stronger requirement for consistency than the usual condition stated in (1.14). In particular, (2.3) is responsible for the upwind nature of the methods under consideration as stated in (2.9) below. Section III will give examples of numerical flux functions under the form (2.7) that are Lipschitz continuous and consistent with  $f$  in the sense of (1.14) but such that (2.3) is not satisfied.

2.3- For a given  $s$  in  $[0, 1]$ , neither  $B^+(\cdot, s)$  nor  $B^-(\cdot, s)$  are assumed to be Jacobian matrices. Therefore, the vector-valued integrals in (2.5) and (2.7) are in general path dependent.

According to Definition 1, when all the signal velocities associated with the numerical flux  $f(u_L, u_R)$  are nonnegative, i.e.  $B^-$  identically vanishes along the path  $\Phi(\cdot; u_L, u_R)$ ; then from the compatibility condition (2.3), we easily deduce that the numerical flux-function (2.7) must reduce to :

$$f(u_L, u_R) = \mathbf{f}(u_L). \quad (2.9a)$$

On the other hand, when all the signal velocities associated with the numerical flux  $f(u_L, u_R)$  are nonpositive; i.e.  $B^+$  identically vanishes along the path  $\Phi(\cdot; u_L, u_R)$  then

$$f(u_L, u_R) = \mathbf{f}(u_R). \quad (2.9b)$$

As a consequence, we state

### Corollary 1

*Any given  $(\Phi; B^\pm)$  scheme is a 3-point upstream method (1.9)-(1.12) in the sense of (2.9).*

The definition we propose meets the usual notion of upwind bias. Conversely, we do not claim that any given consistent and conservative 3-point upstream method in the rather vague sense of (2.9) belongs to the  $(\Phi; B^\pm)$  formulation. However, the examples given below show that our formalism is quite natural and indeed sufficiently general to recover the classical approximate Riemann solvers (such as flux difference splitting methods) as well as the flux vector splitting methods. Moreover, this formalism enables a precise distinction to be made between these two approaches. Roughly speaking, we shall see that FVS methods completely ignore the concept of waves (or field by field) decomposition which is by contrast the starting point of all the FDS methods.

### Proof of the Assertion

Let us first introduce the following two functions  $f_L$  and  $f_R : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}^p$ , setting for any  $u_L$  and  $u_R$  in  $\mathcal{U}$

$$f_L(u_L, u_R) = \mathbf{f}(u_L) + \int_0^1 B^-(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds, \quad (2.10a)$$

$$f_R(u_L, u_R) = \mathbf{f}(u_R) - \int_0^1 B^+(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds. \quad (2.10b)$$

Equipped with (2.10a) and (2.10b), the averaging (2.5a) and (2.5b) can be rewritten under the form

$$\tilde{u}_L(u_L, u_R) = u_L - \frac{2h}{h_x} (f_L(u_L, u_R) - \mathbf{f}(u_L)), \quad (2.11a)$$

$$\tilde{u}_R(u_L, u_R) = u_R - \frac{2h}{h_x} (\mathbf{f}(u_R) - f_R(u_L, u_R)). \quad (2.11b)$$

In view of the compatibility condition (2.3), we have

$$\begin{aligned} f_L(u_L, u_R) &= f_R(u_L, u_R) = \frac{1}{2} (f_L(u_L, u_R) + f_R(u_L, u_R)) \\ &\equiv f(u_L, u_R), \quad \forall u_L, u_R \in \mathcal{U}, \end{aligned}$$

so that it makes sense to use (2.7) as a numerical flux-function and owing to (2.11), definition (2.6) clearly yields a 3-point scheme in conservation form.

We now turn to proving the local Lipschitz continuity of the flux-function (2.7). Given a bounded set  $K$  of  $\mathcal{U}$ , then for any  $(u_L, u_R)$  and  $(v_L, v_R)$  in  $K$  we write

$$\begin{aligned} f(v_L, v_R) - f(u_L, u_R) &= \frac{1}{2} (\mathbf{f}(v_R) - \mathbf{f}(u_R)) + \frac{1}{2} (\mathbf{f}(v_L) - \mathbf{f}(u_L)) \\ &- \frac{1}{2} \int_0^1 \left\{ |B|(\Phi(s; v_L, v_R), s) \frac{\partial \Phi(s; v_L, v_R)}{\partial s} - |B|(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right\} ds. \end{aligned} \quad (2.12)$$

Note that the above integral can be split into three terms according to

$$\begin{aligned} I_1 + I_2 + I_3 &= + \frac{1}{2} \int_0^1 |B|(\Phi(s; v_L, v_R), s) \left\{ \frac{\partial \Phi(s; v_L, v_R)}{\partial s} - \frac{\partial \Phi(s; u_L, v_R)}{\partial s} \right\} ds \\ &+ \frac{1}{2} \int_0^1 \left\{ |B|(\Phi(s; v_L, v_R), s) - |B|(\Phi(s; u_L, u_R), s) \right\} \frac{\partial \Phi(s; u_L, v_R)}{\partial s} ds \\ &- \frac{1}{2} \int_0^1 |B|(\Phi(s; u_L, u_R), s) \left\{ \frac{\partial \Phi(s; u_L, u_R)}{\partial s} - \frac{\partial \Phi(s; u_L, v_R)}{\partial s} \right\} ds. \end{aligned}$$

Using successively the smoothness condition (2.1c) satisfied by  $\Phi$  and (2.2b), we obtain the following crude estimates

$$\begin{aligned} |I_1| &\leq C (\|B^+\| + \|B^-\|) \{ |(v_R - v_L) - (v_R - u_L)| \} \\ &\leq C |v_L - u_L|, \end{aligned}$$

and similarly,

$$|I_3| \leq C |v_R - u_R|.$$

Moreover, in view of (2.1d) and assumption (2.2c), we deduce the following bound for  $I_2$

$$\begin{aligned} |I_2| &\leq C \max_{s \in [0,1]} \{ |\Phi(s; v_L, v_R) - \Phi(s; u_L, u_R)| \} |v_R - u_L|, \\ &\leq C (|u_L| + |v_R|) \max_{s \in [0,1]} \{ |\Phi(s; v_L, v_R) - \Phi(s; u_L, u_R)| \}. \end{aligned}$$

Note that for  $s \in [0, 1]$  :

$$\begin{aligned} |\Phi(s; v_L, v_R) - \Phi(s; u_L, u_R)| &\leq |(\Phi(s; v_L, v_R) - v_L) - (\Phi(s; u_L, u_R) - u_L)| + |v_L - u_L| \\ &\leq \left| \int_0^s \left\{ \frac{\partial \Phi(t; v_L, v_R)}{\partial t} - \frac{\partial \Phi(t; u_L, u_R)}{\partial t} \right\} dt \right| + |v_L - u_L| \\ &\leq s C \{ |(v_R - v_L) - (u_R - u_L)| \} + |v_L - u_L| \\ &\leq C \{ |v_R - u_R| + |v_L - u_L| \}, \end{aligned}$$

where we have used (2.1a) and (2.1c). Therefore, there exists a constant  $C$  depending on  $K$  such that

$$|f(v_L, v_R) - f(u_L, u_R)| \leq C \{ |v_R - u_R| + |v_L - u_L| \}.$$

This gives the required result.

We now conclude by proving the consistency of the numerical flux (2.7) with the exact flux-function  $\mathbf{f}$ . Using the consistency condition (2.1b) concerning the path  $\Phi$ , we have for any  $u$  in  $\mathcal{U}$

$$\frac{\partial \Phi(s; u, u)}{\partial s} = 0, \quad \text{a.e. } s \text{ in } [0, 1], \quad (2.13)$$

since (2.1b) does not depend on  $s$ , so that  $f(u, u) = \mathbf{f}(u)$ . This concludes the proof.  $\square$

We now show how the classical FDS and FVS methods can be re-expressed using the formalism of Definition 1.

## II.1 Flux Difference Splitting (FDS) Methods

Flux difference splitting methods, also referred to as approximate Riemann solvers, achieve upstream biasing on the basis of a given field by field decomposition. Such a decomposition intends to restore, but via a much simpler structure, some of the important features of the Godunov exact Riemann solver. Essential among the features required to be satisfied is the property of conservation (2.8b).

A full hierarchy of approximate Riemann solvers exists, ranging from the simplest one with only two waves and one intermediate state [22] to the most sophisticated ones

involving a complete set of  $p$  waves separating  $(p - 1)$  intermediate states (for instance Roe [45], Collela-Glaz [4]). The Osher-Solomon method [35] is not strictly speaking an approximate Riemann solver but nevertheless clearly falls by construction into our setting. We show below how to re-express any given approximate Riemann solver in terms of the formalism we have introduced in Definition 1. For the sake of clarity, the derivation we propose is performed at a formal level although it can be rigorously justified using the notion of graph completion introduced by DalMaso, LeFloch and Murat [9]. For this construction to be valid, we ask the physical space  $\mathcal{U}$  to be a convex set. Moreover, approximate Riemann solutions are assumed to be made of only simple waves as for the exact solution (cf. assumptions (1.5a) and (1.5b)). These assumptions indeed simplify the derivation. Applications of this formal construction to classical approximate Riemann solvers are postponed to Appendix 1.

By definition, a FDS method provides us with a self similar function  $w(\cdot; u_L, u_R) : \xi \in \mathcal{R} \rightarrow w(\xi; u_L, u_R) \in \mathcal{U}$  where  $\xi = x/t$ , such that  $w(-\infty; u_L, u_R) = u_L$ ,  $w(+\infty; u_L, u_R) = u_R$ . Generally speaking, such a function, referred as to an approximate Riemann solution, exists provided that  $u_L$  and  $u_R$  are close enough. We shall assume that the approximate Riemann solutions under consideration are piecewise smooth except on a finite set of points  $(\xi_d)_{1 \leq d \leq d_{max}}$  where  $w(\cdot; u_L, u_R)$  admits left and right limits  $w(\xi_d^\pm; u_L, u_R)$ . This assumption is satisfied by the known FDS methods. Therefore, the total variation of such solutions is bounded, i.e.

$$TV_{-\infty}^{+\infty}(w(\cdot; u_L, u_R)) = \int_{-\infty}^{+\infty} \left| \frac{dw}{dy}(y; u_L, u_R) \right| dy < C, \quad (2.14)$$

where the above gradient is understood as a vector-valued bounded measure acting on  $\mathcal{R}$ . For convenience, we shall adopt the strict inequality in the CFL condition (2.4), that is to consider times  $t$ ,  $t < h$  during which the waves are strictly contained in the rectangle  $] -h_x/2, h_x/2[ \times ]0, h[$ . It is thus equivalent to restricting our attention to  $\xi \in ] -\xi_0, +\xi_0[$  where

$$\xi_0 = \frac{h_x}{2h}. \quad (2.15)$$

With this restriction, we clearly have

$$TV_{-\infty}^{+\infty}(w(\cdot; u_L, u_R)) = TV_{-\xi_0}^{+\xi_0}(w(\cdot; u_L, u_R))$$

since for all  $\xi \in \mathcal{R}^+$ ;  $\xi \geq \xi_0$ , we have  $w(-\xi; u_L, u_R) = u_L$  and  $w(+\xi; u_L, u_R) = u_R$ .

For an approximate Riemann solution to be relevant, we seek  $w(\cdot; u_L, u_R)$  to satisfy the averaging consistent with the exact Godunov Riemann solver (cf. (2.8b)),

$$\frac{1}{h_x} \int_{-h_x/2}^{h_x/2} w\left(\frac{x}{h}; u_L, u_R\right) dx = \frac{1}{2}(u_L + u_R) - \frac{h}{h_x}(\mathbf{f}(u_R) - \mathbf{f}(u_L)). \quad (2.16a)$$

In view of (2.15), this requirement reduces to

$$\int_{-\xi_0}^{+\xi_0} w(\xi; u_L, u_R) d\xi = \xi_0(u_L + u_R) - (\mathbf{f}(u_R) - \mathbf{f}(u_L)). \quad (2.16b)$$

Now, let  $\mathcal{C}(w) = ] - \xi_0, +\xi_0[ \setminus \{\xi_d\}_{1 \leq d \leq d_{\max}}$  denote the open set of points where  $w(\cdot; u_L, u_R)$  is continuous. If we notice that we have :

$$\begin{aligned} \int_{\mathcal{C}(w)} \xi \frac{dw}{d\xi}(\xi; u_L, u_R) d\xi + \sum_{d=1}^{d_{\max}} \xi_d (w(\xi_d^+; u_L, u_R) - w(\xi_d^-; u_L, u_R)) = \\ \xi_0(u_L + u_R) - \int_{-\xi_0}^{+\xi_0} w(\xi; u_L, u_R) d\xi, \end{aligned} \quad (2.17)$$

then the relation (2.16) can be re-expressed as follows

$$\begin{aligned} \int_{\mathcal{C}(w)} \xi \frac{dw}{d\xi}(\xi; u_L, u_R) d\xi + \sum_{d=1}^{d_{\max}} \xi_d (w(\xi_d^+; u_L, u_R) - w(\xi_d^-; u_L, u_R)) = \\ + (\mathbf{f}(u_R) - \mathbf{f}(u_L)). \end{aligned} \quad (2.18)$$

The identity (2.18) plays in fact a central role in the derivation we perform below. Roughly speaking; when considering (2.18), the compatibility condition (2.3) indeed suggests that the physical states  $w(\xi; u_L, u_R)$ ,  $\xi \in ] - \xi_0, +\xi_0[$ , may serve as a natural variable to define the states involved in the path  $\Phi(\cdot; u_L, u_R)$  and that  $\xi$  may serve, depending on its sign, to define either  $B^+$  or  $B^-$ . Indeed, notice that wherever the self similar function  $w(\cdot; u_L, u_R)$  is smooth and non-constant in a neighborhood of  $\xi$ , then  $\xi$  returns the velocity at which the state  $w(\xi; u_L, u_R)$  propagates. Moreover,  $\xi_d$ ,  $d \in \{1, \dots, d_{\max}\}$ , does coincide with the velocity of the  $d$ th discontinuity separating  $w(\xi_d^-; u_L, u_R)$  from  $w(\xi_d^+; u_L, u_R)$ . Our purpose in the sequel is to develop a precise construction to the above assertion in particular regarding the treatment of the discontinuities. A geometrical picture depicted in Fig. 1 is useful to understand the following mathematical construction, in which Fig. 1a is a set of physical states displaying various wave structures.

To facilitate our construction of the framework, it is convenient to introduce the following notation:

$$TV_{-\xi_0}^{\xi}(w) = \int_{-\xi_0}^{\xi} \left| \frac{dw}{dy}(y; u_L, u_R) \right| dy, \quad \forall \xi \in ] - \xi_0, +\xi_0[. \quad (2.19)$$

Let us consider the mapping  $\tilde{X} : \xi \in \mathcal{R} \rightarrow \tilde{X}(\xi) \in [0, 1]$  given by

$$\tilde{X}(\xi) = \frac{1}{TV_{-\xi_0}^{+\xi_0}(w)} TV_{-\xi_0}^{\xi}(w(\cdot; u_L, u_R)). \quad (2.20)$$

Note that (2.20) defines a left continuous increasing function that is discontinuous across each discontinuity shown in Fig. 1b. To make use of this discontinuous data for our following purpose, we define a pair of functions  $\theta^\pm : \xi \in [-\xi_0, +\xi_0[ \rightarrow \theta^\pm(\xi) \in ]-\xi_0, +\xi_0[$

$$\theta^+(\xi) = \min_{\mu \in ]-\xi_0, +\xi_0[} \{ \mu; \tilde{X}(\mu) > \tilde{X}(\xi) \}, \quad (2.21a)$$

and

$$\theta^-(\xi) = \min_{\mu \in ]-\xi_0, +\xi_0[} \{ \mu; \tilde{X}(\mu) = \tilde{X}(\xi) \}. \quad (2.21b)$$

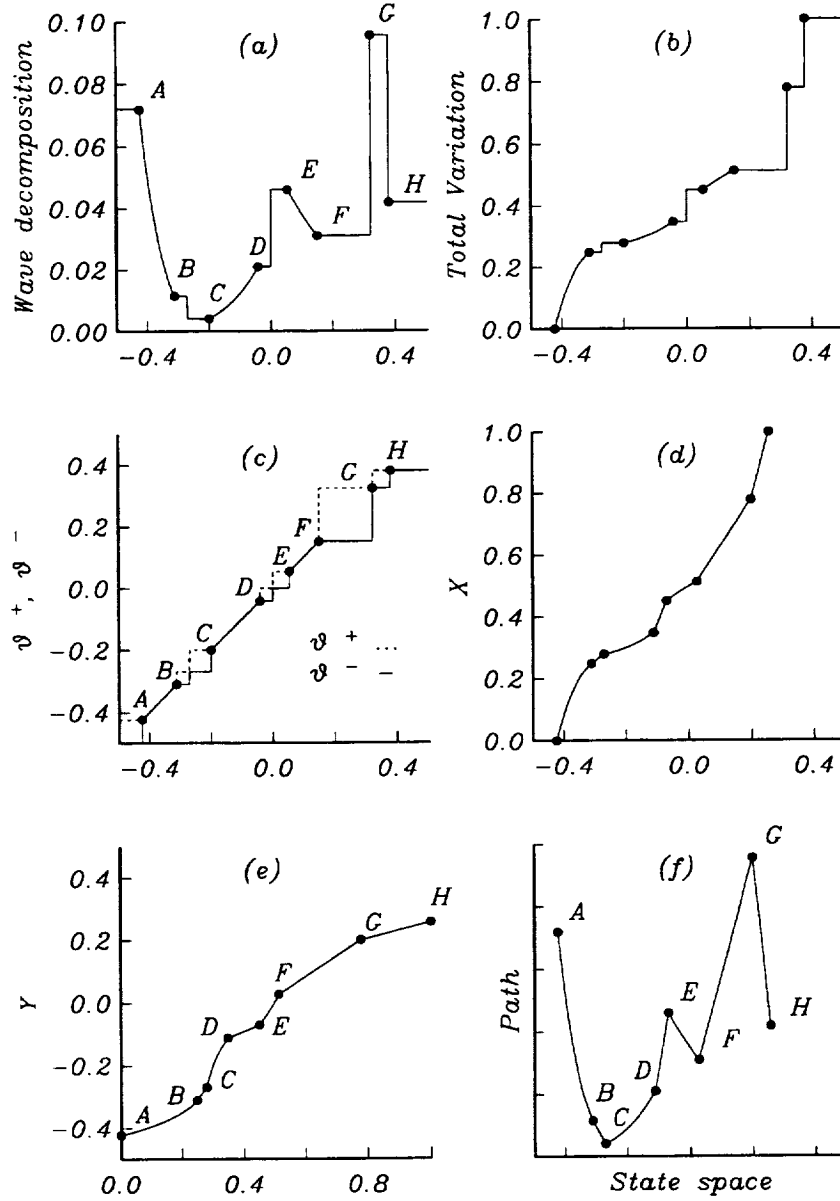


Figure 1. Phase diagram: (a) Wave distribution, (b) Total variation, (c)  $\theta^\pm(\xi)$ , (d)  $\chi(\mu)$ , (e)  $Y(s)$ , and (f) Path  $\Phi$ .



Note that  $\xi \rightarrow |\theta^+(\xi) - \theta^-(\xi)|$  yields a left continuous step function, as depicted in Fig. 1c. Indeed, wherever the approximate Riemann solution  $w(\cdot; u_L, u_R)$  is smooth and non-constant in a neighborhood of  $\xi$ , then  $|\theta^+(\xi) - \theta^-(\xi)|$  is zero since  $\theta^+(\xi) = \theta^-(\xi) = \xi$ . Otherwise, this function gives the length  $|I_k|$  of the interval  $I_k$  for a given  $k$  such that  $\xi \in I_k$  and

$$w(\mu; u_L, u_R) = u_k, \quad \forall \mu \in I_k. \quad (2.22)$$

Here  $u_k$  denotes the  $k^{th}$  constant state of  $w(\cdot; u_L, u_R)$ . For such values of  $\xi \in I_k$ , we indeed have

$$\theta^-(\xi) = \min_{\xi \in I_k} \xi, \quad \theta^+(\xi) = \max_{\xi \in I_k} \xi. \quad (2.23)$$

For (2.22) to be valid, it is implicitly assumed that if there exist two adjacent intervals  $I_k$  and  $I_{k+1}$ , i.e if the constant states  $u_k$  and  $u_{k+1}$  are separated by a discontinuity, then the associated lengths  $|I_k|$  and  $|I_{k+1}|$  are distinct. In what follows, this assumption is supposed to be met. (If not, we would have to modify  $|\theta^+ - \theta^-|$  using a relevant  $\xi$ -dependent factor to achieve a discontinuity from  $I_k$  to  $I_{k+1}$ .)

Let us denote by  $\xi_l$  the location of the  $l^{th}$  discontinuity of the function  $\xi \in ]-\xi_0, \xi_0[ \rightarrow |\theta^+ - \theta^-|(\xi)$ . For completeness, we introduce  $\xi_{l_{max}+1} = +\xi_0$  in such a way that the set  $\cup_0^{l_{max}} [\xi_l, \xi_{l+1}[$  yields a partition of  $[-\xi_0, +\xi_0[$ . In the sequel,  $J_l$  denotes the interval  $[\xi_l, \xi_{l+1}[$ . For our purpose, we now want to remove from this partition any interval  $J_l$  such that  $|\theta^+ - \theta^-|$  is zero within  $J_{l+1}$ .

In this way, we construct an interval  $S$  of  $[-\xi_0, +\xi_0[$  using the procedure

$$S_0 = \begin{cases} J_0, & \text{if } |\theta^+ - \theta^-| \text{ is not zero within } J_1 \\ \emptyset, & \text{otherwise;} \end{cases} \quad (2.24a)$$

$$S_1 = \begin{cases} J_1, & \text{if } |\theta^+ - \theta^-| \text{ is not zero within } J_2 \\ \emptyset, & \text{otherwise.} \end{cases} \quad (2.24b)$$

Note that  $S_0$  and  $S_1$  cannot be simultaneously empty sets. Let us denote by  $\mu_0$  the minimal value of  $\xi$  within  $S_0 \cup S_1$ . We then complete our construction for  $k \leq l_{max}$  as follows

$$S_k = [\mu_k, \mu_{k+1}[ , \quad \mu_k = \mu_0 + \sum_{l=0}^{k-1} |S_l|, \quad \mu_{k+1} = \mu_k + |S_k| \quad (2.24c)$$

$$\text{where } |S_k| = \begin{cases} |J_k|, & \text{if } |\theta^+ - \theta^-| \text{ is not zero within } J_{k+1} \\ 0, & \text{otherwise.} \end{cases} \quad (2.24d)$$

Note that when  $|S_k| \neq 0$ , i.e.  $S_k \neq \emptyset$ , then

$$\xi - \xi_k = \mu - \mu_k, \quad \forall \mu \in S_k, \forall \xi \in J_k; \quad (2.25)$$

so that  $S_k$  can be deduced from  $J_k$  by a suitable translation. Finally, we define

$$S = \cup_0^{l_{max}} S_k. \quad (2.26)$$

We now reconstruct a new function  $X : \mu \in S \rightarrow X(\mu) \in [0, 1]$  using the following procedure. Let us consider a given nonempty subset  $S_k \in S$ , then for  $\mu \in S_k$  and its related  $\xi \in J_k$  given by (2.25), we define

$$X(\mu) = \begin{cases} \tilde{X}(\xi), & \text{if } \theta^+(\xi) - \theta^-(\xi) = 0 \\ \tilde{X}(\theta^-(\xi)) + \mathcal{H}(\mu)(\tilde{X}(\theta^+(\xi)) - \tilde{X}(\theta^-(\xi))), & \text{otherwise.} \end{cases} \quad (2.27a)$$

Here, we have set for  $\mu \in S_k$

$$\mathcal{H}(\mu) = \frac{\mu - \theta^-(\xi)}{\theta^+(\xi) - \theta^-(\xi)}, \quad \text{where } \xi - \xi_k = \mu - \mu_k. \quad (2.27b)$$

This step does nothing but connecting the jumps of  $w(\cdot; u_L, u_R)$  with a straight line, as seen in Fig. 1d. As a consequence,  $X : \mu \in S \rightarrow X(\mu) \in [0, 1]$  is a continuous strictly increasing function. It therefore makes sense to consider its inverse function  $Y : s \in [0, 1] \rightarrow Y(s) \in S$  such that for a given  $s \in [0, 1]$ ,  $Y(s) = \mu$  where  $\mu$  is the unique solution in  $S$  of  $s = X(\mu)$ .  $Y$  is depicted in Fig. 1e.

Equipped with this function  $Y$ , we finally construct a path  $\Phi$  as follows for  $s \in [0, 1]$ :

$$\Phi(s; u_L, u_R) = \begin{cases} w(\xi; u_L, u_R), & \text{if } \theta^+(\xi) - \theta^-(\xi) = 0 \\ w(\theta^-(\xi); u_L, u_R) + \mathcal{H}(\xi)(w(\theta^+(\xi); u_L, u_R) - w(\theta^-(\xi); u_L, u_R)); \end{cases} \quad (2.28a)$$

where for  $s \in [0, 1]$ ,  $\xi$  is the solution :

$$\exists k \in \{1, l_{max}\}; \quad \begin{cases} \xi - \xi_k = \mu - \mu_k \\ Y(s) = \mu. \end{cases} \quad (2.28b)$$

Notice that according to the above construction, a constant intermediate state  $u_k$  separated by two waves from the decomposition under consideration reduces to the point  $u_k \in \Phi(\cdot, u_L, u_R)$  while two states  $u_k$  and  $u_{k+1}$  separated by a discontinuity (either a shock or a contact) are connected along  $\Phi(\cdot; u_L, u_R)$  by a straight line (see Fig. 1f). Indeed for

$S_k \neq \emptyset$  such that  $|\theta^+ - \theta^-|$  is not zero within  $J_k$ , there exists an index  $d \in \{1, \dots, d_{\max}\}$  such that

$$\begin{cases} w(\theta^-(\xi); u_L, u_R) = w(\xi_d^-; u_L, u_R) \\ w(\theta^+(\xi); u_L, u_R) = w(\xi_d^+; u_L, u_R), & \xi - \xi_k = \mu - \mu_k, \quad \forall \mu \in S_k, \\ \theta^+(\xi) = \xi_d \end{cases} \quad (2.29)$$

where  $\xi_d$  is the location of the  $d$ th discontinuity of  $w(\cdot; u_L, u_R)$ .

The case of the rarefaction waves, that is of the states  $w(\xi; u_L, u_R)$  with  $\xi \in ]-\xi_0, \xi_0[$  such that  $\theta^+(\xi) = \theta^-(\xi) = \xi$ , is straightforward.

Now turning to the matrices  $B^\pm$ , we set for  $s \in [0, 1]$

$$B^+(\Phi(s; u_L, u_R), s) = \frac{1}{2}(\theta^+(\xi) + |\theta^+(\xi)|) \cdot Id_{\mathcal{R}^p}, \quad (2.30a)$$

$$B^-(\Phi(s; u_L, u_R), s) = \frac{1}{2}(\theta^+(\xi) - |\theta^+(\xi)|) \cdot Id_{\mathcal{R}^p}, \quad (2.30b)$$

where  $\xi$  is the unique solution of the problem stated in (2.28b). Note that the two matrix-valued functions  $B^\pm$  defined in (2.30) trivially satisfy the requirement (2.3) for upwind biasing.

Furthermore, in view of (2.18) satisfied by the approximate solution  $w(\cdot; u_L, u_R)$ , we clearly deduce from (2.29) and definitions (2.28) and (2.30) the compatibility expression:

$$\begin{aligned} \int_0^1 (B^+ + B^-)(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds &= \sum_{S_k; |\theta^+ - \theta^-| = 0} \int_{\xi_k}^{\xi_{k+1}} \xi \frac{dw}{d\xi}(\xi; u_L, u_R) d\xi \\ &+ \sum_{d=1}^{d_{\max}} \xi_d (w(\xi_d^+; u_L, u_R) - w(\xi_d^-; u_L, u_R)) = \mathbf{f}(u_R) - \mathbf{f}(u_L). \end{aligned}$$

From the construction performed above, the fixed family of paths  $\Phi$  associated with a given approximate Riemann solver is in fact made of the physical states involved in the wave decomposition in the zone of smoothness of the approximate solution  $w(\cdot; u_L, u_R)$  while the matrix-valued functions  $B^\pm$  return the associated wave velocities. Finally, a straight line is used to connect in the phase space  $\mathcal{U}$  two states separated by a discontinuity and  $B^\pm$  are defined thanks to the associated discontinuity velocity. We emphasize that the family of paths  $\Phi$  and its associated  $B^\pm$  functions completely depend on the approximate Riemann solution under consideration, that is on the given FDS method itself.

In Section III, we shall need to make a partition of the path  $\Phi(\cdot; u_L, u_R)$  (2.28) associated with a given FDS method, into  $l_{\max}$  nonoverlapping subpaths  $\tilde{\Phi}_k(\cdot; u_L, u_R)$  with

$k \in \{1, \dots, l_{max}\}$ . The  $k$ th subpath  $\tilde{\Phi}_k(\cdot; u_L, u_R) : s \in [s_k, s_{k+1}] \rightarrow \tilde{\Phi}_k(s; u_L, u_R) \in \mathcal{U}$  is associated with the  $k$ th subinterval  $S_k = [\mu_k, \mu_{k+1}[$  of  $S$  in the following way. Here  $s_k$  and  $s_{k+1}$  in  $[0, 1]$  are chosen to be

$$\Phi(s_k; u_L, u_R) = w(\xi_k; u_L, u_R); \quad \Phi(s_{k+1}; u_L, u_R) = w(\xi_{k+1}; u_L, u_R). \quad (2.31)$$

We then define

$$\tilde{\Phi}_k(s; u_L, u_R) = \Phi(s; u_L, u_R), \quad s \in [s_k, s_{k+1}], \quad (2.32)$$

to rewrite

$$\Phi(\cdot; u_L, u_R) = \bigcup_{k=1}^{l_{max}} \tilde{\Phi}_k(\cdot; u_L, u_R). \quad (2.33)$$

For convenience in the notation, we shall make use in the sequel of

$$\Phi_k(s; u_L, u_R) = \tilde{\Phi}_k(s_k + s(s_{k+1} - s_k); u_L, u_R), \quad \forall s \in [0, 1], \quad (2.34a)$$

so that now

$$\Phi_k(0; u_L, u_R) = w(\xi_k; u_L, u_R), \quad \Phi_k(1; u_L, u_R) = w(\xi_{k+1}; u_L, u_R). \quad (2.34b)$$

Besides the formal similarity of their expression, the revisited approximate Riemann solvers are by construction path dependent. Put in other words, their associated numerical flux-function entirely depends on the waves (or field by field) decomposition under consideration. In turn, the design of such a decomposition is not only responsible for the merits but also for the deficiencies of the resulting approximate Riemann solver.

Besides robustness, one of the critical issues in the derivation is how to achieve high resolution of stationary discontinuities while at the same time ensuring enough nonlinear dissipation to select the physically relevant discontinuous solutions. The difficulty in achieving this goal comes from the fact that entropy violation only occurs for stationary or near stationary discontinuities associated with the genuinely nonlinear fields.

The current prominent approximate Riemann solvers have been designed to perfectly resolve a stationary contact discontinuity ([4], [35], [45]) and to yield a sharp steady discrete shock [4] or to exactly capture a shock ([35], [45]). But associated with this desirable accuracy, their nature of approximation results in insufficient robustness in calculating some practical problems. This is in particular true for the Roe and Collela-Glaz methods that may fail in the capture of large expansion waves, both admitting nonphysical shocks as stable solutions. A lack of robustness has been reported for the Osher scheme in the capture of strong detached shocks [24].

Fixes have been introduced to cure some of the discussed failures [12], [21], [24], [46], but at the expense of losing accuracy. These extra mechanisms need to be properly tuned up either to be efficient or to restore locally some desirable accuracy that was lost by the uncorrected scheme. This is, in particular, the case of the widely used entropy fix proposed by Harten [21] which is applied in practice to all the fields of the Roe decomposition in order to increase the overall robustness of the method. This procedure in turn destroys the high resolution of the Roe scheme for stationary discontinuities. We stress that the exact capture of a contact discontinuity at rest is a crucial requirement for the correct resolution of boundary layers [50], [14] (see also section V).

We now turn to re-express FVS methods in the framework of Definition 1.

## II.2 Flux Vector Splitting (FVS) Methods.

In this section, we study a different and widely used approach for implementing upwind biasing, namely the FVS methods. The statements given below will underline the deep differences from the FDS methods we have just revisited in our setting. We first state

### Definition 2

*Let  $(f^+, f^-)$  be a pair of continuous and piecewise differentiable functions  $f^\pm : \mathcal{U} \rightarrow \mathcal{R}^p$ , required to satisfy the following two conditions*

$$f^+(u) + f^-(u) = \mathbf{f}(u), \quad u \in \mathcal{U}, \quad (2.35)$$

*and in any region of smoothness*

$$\begin{aligned} B^+(u) \equiv df^+(u) \text{ (respectively } B^-(u) \equiv df^-(u)) \text{ is diagonalizable} \\ \text{with real and nonnegative (resp. real and nonpositive) eigenvalues.} \end{aligned} \quad (2.36)$$

*For any  $u_L$  and  $u_R$  in  $\mathcal{U}$ , the FVS method defines the numerical flux-function associated with the pair  $(f^+, f^-)$  to be*

$$f(u_L, u_R) = f^+(u_L) + f^-(u_R). \quad (2.37)$$

### Remarks

2.4- The exact flux-function  $\mathbf{f}$  needs not be a homogeneous function of  $u$  of degree one for a consistent FVS method to exist (in the sense of (2.35)), as pointed out in [22].

2.5- In view of the requirement (2.36), the definition of the FVS methods we give is different from the one originally given in [47] and [22]. However, we stress that all the

FVS methods that have been brought to our attention, satisfy the condition (2.36) (see Lerat [29] for the Steger-Warming splitting [47]). Moreover, the Van Leer method satisfies by construction the above definition while by contrast it does not fall into the framework for splitting methods proposed in [47] (see Hirsh [23]). The implications we give of the Definition 2 will be assessed in what follows (see, in particular, Propositions 2).

As an immediate consequence of Definition 2, we state

**Corollary 2**

*Let be given a general FVS method associated with the pair  $(f^+, f^-)$ . This FVS method is a  $(\Phi; B^\pm)$  method for any given fixed family of path  $\Phi$ . Here,  $B^\pm$  are the Jacobian matrices defined by (2.96). The numerical flux function defined in (2.7) can be re-expressed as*

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \int_0^1 (B^+ - B^-)(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds \right), \quad (2.38)$$

*or under the equivalent form*

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{-1, +1} \text{sign}(\pm 1) (f^\pm(u_R) - f^\pm(u_L)) \right). \quad (2.39)$$

FVS methods therefore fall into our framework for upwind biasing but according to our formalism, FVS methods are *path independent*. We stress that by this strong property, a general FVS method is in complete opposition with the basically path dependent FDS methods. This deep distinction, as provided by our formalism, will be given a physical interpretation in the next section devoted to the Euler equations.

Remark

2.6- A formal explanation of the path independence can be given on the basis of the following two systems of conservation laws :

$$\frac{\partial u}{\partial t} + \frac{\partial f^-(u)}{\partial x} = 0,$$

$$\frac{\partial u}{\partial t} + \frac{\partial f^+(u)}{\partial x} = 0,$$

associated with the following initial data

$$u(0, x) = \begin{cases} u_L(x), & \text{if } x < 0, \\ u_R(x), & \text{if } x \geq 0. \end{cases}$$

We shall assume, although it is not clear with no further assumption on  $f^+$  and  $f^-$ , that both the Riemann problems above admit solutions, repectively denoted by  $w(., u_L, u_R; f^-)$  and  $w(., u_L, u_R; f^+)$ . Now, in view of Remark 2.1, the averaging (2.5a) applied to  $w(., u_L, u_R; f^-)$  yields at least formally

$$\tilde{u}_L(u_L, u_R) = u_L - \frac{2h}{h_x}(f^-(u_R) - f^-(u_L)),$$

since under the requirement (2.36), no wave can propagate outside of the rectangle  $(-h_x/2, 0) \times (0, h)$ . Conversely and using again (2.36), we can write for the averaging (2.5b) but with  $w(., u_L, u_R; f^+)$

$$\tilde{u}_R(u_L, u_R) = u_R - \frac{2h}{h_x}(f^+(u_R) - f^+(u_L)).$$

Notice that the above averaging can be rewritten under the expected forms

$$\begin{aligned}\tilde{u}_L(u_L, u_R) &= u_L - \frac{2h}{h_x}((f^-(u_R) + f^+(u_L)) - \mathbf{f}(u_L)), \\ \tilde{u}_R(u_L, u_R) &= u_R - \frac{2h}{h_x}(\mathbf{f}(u_R) - (f^-(u_R) + f^+(u_L))),\end{aligned}$$

where we have used the consistency condition (2.35).

The averaging procedures performed above have made no use of the precise wave patterns involved in  $w(., u_L, u_R; f^\pm)$ . The averaged solutions would have been the same for any given other wave decompositions, provided that the waves under consideration are only propagating in their respective relevant rectangle (see requirement (2.36)). The path independence property stated in Corollary 2, may be understood in that way.

To conclude the present remark, we note that FVS methods can be given at least formally a Godunov-type interpretation (but in sharp contrast with FDS methods), under which they make use of two distinct sets of Riemann problem for evaluating the same numerical flux (cf. in particular Remark 2.8).

### *Proof of Corollary 2*

The path independence property clearly comes from the Jacobian nature of the matrices  $B^\pm$ . Indeed, we have for any given fixed family of paths in  $\mathcal{U}$

$$\int_0^1 B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = f^+(u_R) - f^+(u_L), \quad (2.40a)$$

$$\int_0^1 B^-(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = f^-(u_R) - f^-(u_L). \quad (2.40b)$$

Therefore, substituting (2.40a) and (2.40b) in (2.38) clearly gives the expected flux formula (2.39). Now using the consistency condition (2.35), let us write

$$\mathbf{f}(u_L) + \mathbf{f}(u_R) = (f^+(u_L) + f^-(u_L)) + (f^+(u_R) + f^-(u_R)). \quad (2.41)$$

Formula (2.37) directly follows by substitution in (2.39).

Moreover, by differentiating the consistency relation (2.35), we deduce the following identity

$$B^+(u) + B^-(u) = A(u), \quad \text{a.e. } u \in \mathcal{U}, \quad (2.42)$$

where  $A$  is the Jacobian matrix associated with the flux function  $\mathbf{f}$ . So the compatibility condition (2.3) stated in Definition 1 is satisfied for any given path  $\Phi$ . This concludes the proof.  $\square$

### Remarks

2.7- Although the identity  $B^+(u) + B^-(u) = A(u)$  holds true for almost every  $u$  in  $\mathcal{U}$ , we do not have in general  $B^+(u) = A^+(u)$  and (therefore)  $B^-(u) = A^-(u)$  for all  $u$  in  $\mathcal{U}$  (see Steger and Warming [47]). If these equalities were met for all  $u$ , the associated FVS method would be necessarily path dependent since  $A^\pm(u)$  are not Jacobian matrices, except precisely when all the eigenvalues of  $A(u)$  keep the same sign. Notice that even for such  $u$ , the equality  $B^\pm(u) = A^\pm(u)$  is not necessarily satisfied. This is, in particular, the case of the  $\Gamma$ -Boltzmann schemes developed in Appendix 2.

2.8- Generally speaking,  $B^+(u)$  and  $B^-(u)$  do not commute. As a consequence, they do *not* admit the same set of right eigenvectors. In that case, notice that the right eigenvectors  $\{r_k(u)\}_{1 \leq k \leq p}$  of  $A(u)$  do not diagonalize either  $B^+(u)$  or  $B^-(u)$ . This property is indeed responsible for some difficulties in the analysis of general FVS methods.

By contrast with the property stated here, formulae (2.30a) and (2.30b) clearly indicate that the  $B^\pm$  matrices do commute for a given FDS method; furthermore, FDS schemes satisfy an *averaged* consistency condition (2.16) while FVS methods merely obey the *pointwise* consistency condition (2.42). To our knowledge, only the Osher scheme satisfies a similar condition (note that Van Leer [49] has recognized the Osher scheme as a FVS method in the case of a single conservation law; for a closely related result, see Brennier [2])

These observations further underline the fundamental differences between FVS and FDS methods.

By the property of independence with respect to path, we have

### **Proposition 2**

*A general FVS method in the sense of Definition 2 cannot exactly capture a stationary discontinuity.*



*Proof of Proposition 2*

Let us consider two states  $u_L$  and  $u_R$  connected by a single discontinuity at rest, either a shock or a contact discontinuity. In view of the averaging (2.5a) and (2.5b), this wave is exactly captured if and only if

$$\tilde{u}_L(u_L, u_R) = u_L \quad \text{and} \quad \tilde{u}_R(u_L, u_R) = u_R, \quad (2.43)$$

which is equivalent to asking for

$$\int_0^1 B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = 0, \quad (2.44)$$

for all path  $\Phi$  connecting  $u_L$  to  $u_R$ . Taking the scalar product of (2.44) with  $\frac{\partial \Phi(t; u_L, u_R)}{\partial t}$  and integrating over  $t \in [0, 1]$ , we necessarily have

$$\begin{aligned} & \int_0^1 dt \int_0^1 \frac{\partial \Phi(s; u_L, u_R)}{\partial s} B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds + \\ & \int_0^1 \int_0^1 \left( \frac{\partial \Phi(t; u_L, u_R)}{\partial t} - \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right) B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds dt = 0. \end{aligned} \quad (2.45)$$

By assumption, for all  $u$  in  $\mathcal{U}$ ,  $B^+(u)$  only admits nonnegative eigenvalues. So that, we necessarily have

$$\frac{\partial \Phi(s; u_L, u_R)}{\partial s} B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \geq 0, \quad \text{a.e. } s \in [0, 1], \quad (2.46)$$

for any given path  $\Phi$ . Taking advantage of the path independence property, we can always choose a path  $\Phi$  such that

$$\frac{\partial \Phi(s; u_L, u_R)}{\partial s} B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} > 0, \quad s \in I \subset [0, 1], \quad (2.47)$$

where  $I$  denotes a *nonnegligeable* open subset of  $[0, 1]$ . Otherwise,  $B^+$  would be identically zero for almost every  $u$  in  $\mathcal{U}$ ; this would in turn imply, thanks to (2.42), that  $B^- = A$  does admit strictly positive eigenvalues (since  $A$  does), which contradicts with our assumption (2.36). So the strict equality (2.47) holds true for some paths  $\Phi$ . Notice that from a generic smoothness assumption (2.1d) made on  $\Phi$ , for any  $u_L$  and  $u_R$  in  $\mathcal{U}$  we have

$$\left| \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right| = O(|u_R - u_L|), \quad \text{a.e. } s \in [0, 1], \quad (2.48)$$

Therefore, we choose paths  $\Phi$  such that

$$\int_0^1 \frac{\partial \Phi(s; u_L, u_R)}{\partial s} B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds > \delta |u_R - u_L|^2, \quad (2.49)$$

where  $\delta$  is a positive constant independent of  $u_L$  and  $u_R$ . On the other hand, writing for a smooth path, for a.e.  $s \in [0, 1]$  :

$$\left| \frac{\partial \Phi(t; u_L, u_R)}{\partial t} - \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right| \leq \max_{s \in [0, 1]} \left| \frac{\partial^2 \Phi(s; u_L, u_R)}{\partial^2 s} \right| (|u_R - u_L| |s - t|). \quad (2.50)$$

Notice that the normalization has been accounted for in the last parenthesis. We see that if the path under consideration is chosen with the following additional smoothness

$$\left| \frac{\partial^2 \Phi(s; u_L, u_R)}{\partial^2 s} \right| = O(|u_R - u_L|), \quad \text{a.e. } s \in [0, 1], \quad (2.51)$$

we necessarily have

$$\int_0^1 \int_0^1 \left( \frac{\partial \Phi(t; u_L, u_R)}{\partial t} - \frac{\partial \Phi(s; u_L, u_R)}{\partial s} \right) B^+(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds dt = O(|u_R - u_L|^3). \quad (2.52)$$

So that the identity (2.44) is clearly impossible if  $u_L$  and  $u_R$  are chosen close enough while connected by a single discontinuity at rest (notice that this is always possible). This concludes the proof.  $\square$

Proposition 2 applies to states  $u_L$  and  $u_R$  connected by a single shock wave with zero speed or a contact discontinuity at rest as well. This result tells us that by contrast with the FDS methods revisited in section 2.1, a stationary contact discontinuity is necessarily smeared when approximated by a general FVS method. In fact, in the applications, it turns out that such a *linear* wave is unendly smeared, in contrast with a *nonlinear* wave as discussed above. The numerical dissipation underlined here has an important consequence in practical issues when dealing with the numerical approximations to some viscous form of problem (1.1) and (1.2). Its implications have been in particular emphasized by Van Leer et al. [50] for the numerical approximation of the Navier-Stokes equations, thus making a general FVS method an *irrelevant* candidate to this purpose.

The proof given above makes the expression (2.45) play a central role. This expression can be related to the amount of the numerical dissipation inevitably performed by a FVS method (for the relations between the rate of the Entropy dissipation and the numerical viscosity we refer to the work by Coquel and Le Floch [5], [6]). Its quadratic strength is in complete agreement with the results obtained in [5] for numerical methods whose numerical viscosity *never vanishes* (see [6] for the deeply different results obtained in the case of FDS-type Godunov method). According to this observation, the Jacobian nature of

$B^\pm$  combined with the *relevant* assumption (2.36) on the *signs of their eigenvalues* therefore seems to inevitably yield numerical methods with *nonvanishing* numerical viscosity.

Put in other words, any attempt to design pure FVS methods with vanishing viscosity or even with very low diffusivity should result in a *violation of the stability requirements* (2.36); namely  $B^\pm$  would necessarily admit eigenvalues with the wrong sign. The issue is a marginal stability for the method and at least a loss of monotonicity (at *first order* of accuracy) for such low diffusive FVS methods. These conclusions are in complete agreement with the observations by Liou and Steffen [31] and Coirier and Van Leer [3] and make “pure FVS-method a dead-end street” for viscous calculations [3].

Besides this negative result, the robustness achieved by FVS methods strongly suggests that some of their advantageous features should be kept in the derivation of new upwind biased procedures. With respect to the formalism provided by Definition 1, FVS methods are entirely characterized by their associated matrix-valued functions  $B^\pm$  as it is asserted by Corollary 2. As a consequence,  $B^\pm$  provide us with the essence to be kept from the FVS method.

Moreover in order to get high resolution of stationary contact discontinuities while preserving robustness, Proposition 2 clearly suggests that at least extra information specifically regarding the contact discontinuity must be recognized. Since the failure met by the pure FVS methods inherits from their path-independence property, the hybridization we have to perform must result in a path-dependent method. Therefore, this additional and essential step must relate (to a large extent) to a fixed family of paths. As revealed by the framework we have introduced earlier, natural families of paths already exist in the current FDS methods and they can be readily utilized.

All these observations naturally suggest that the tools to be involved are a combination, within the framework of Definition 1, of the two existing approaches for upstreaming. The following section describes a way for hybridizing these tools in order to achieve the requirements that have been put forward so far.

### III. Field by Field Hybrid Upwind Splitting (HUS) Methods

The concept of hybrid upwinding we propose, finds a natural definition in the setting of  $(\Phi; B^\pm)$  schemes we have introduced in the last section. This new approach for upwinding is achieved by combining the main interesting features of two distinctly different concepts. The motivation for such an approach stems from several practical and theoretical reasons outlined in the Introduction. Specifically, we want to derive a new class of upstream methods with the following properties :

- ( P1)- keep the very robustness of flux vector splitting in the capture of genuinely nonlinear waves (strong shock as well as large rarefaction waves),
- ( P2)- preserve the property met by the FVS methods to select only entropy satisfying approximate solutions,  
while at the same time
- ( P3)- offer the decisive ability of some flux difference splitting methods to exactly capture stationary contact discontinuities.

In view of requirement (P3), we shall only consider in the sequel  $(\Phi; B_{FDS}^\pm)$  approximate Riemann solvers that perfectly resolve a contact discontinuity at rest.

As suggested above, HUS methods are derived by making a given FVS scheme path dependent with respect to a family of paths coming from a given FDS method. To this end, we notice that the first two requirements (P1), (P2) only address *genuinely nonlinear fields*, while by contrast, requirement (P3) is only related to *linearly degenerate fields*. As a consequence, for a fixed family of paths  $\Phi$  provided by a given FDS method, we are naturally led to describe each “genuinely nonlinear” subpath with FVS-type  $B_{FVS}^\pm$  Jacobian matrices while keeping unchanged for “linearly degenerate” subpaths the original  $B_{FDS}^\pm$  matrices that are compatible with  $\Phi$ .

With clear notation,  $GNL(\Phi)$  denotes the part of  $\Phi$  made up of the subpaths associated with the genuinely nonlinear fields while  $LD(\Phi)$  stands for the remaining subpaths. For our following purpose, we restrict our attention to considering FDS methods that satisfy the following requirement  $LD(\Phi)$ :

$$\sum_{k; \Phi_k \in LD(\Phi)} \int_0^1 (B_{FDS}^+ + B_{FDS}^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \sum_{k; \Phi_k \in LD(\Phi)} (\mathbf{f}(\Phi_k(1; u_L, u_R)) - \mathbf{f}(\Phi_k(0; u_L, u_R))). \quad (3.1)$$

This we shall refer to in the sequel as a compatibility condition for hybridization. Remark 3.1 will assess such a reference. The validity of (3.1) is investigated below for some of the prominent FDS methods; examples and counter examples are given. We show in remark 3.2 that requirement (3.1) indeed provides us with an essential condition for ensuring upwind biasing for the hybrid methods we now introduce.

**Definition and Proposition 3** *Let be given a general  $(\cdot, B_{FVS}^\pm)$  FVS method and a  $(\Phi, B_{FDS}^\pm)$  FDS method subject to the condition (3.1). A Hybrid Upwind Splitting (HUS) method is defined by the following consistent numerical flux-function :*

$$f(u_L, u_R) = \frac{1}{2} (\mathbf{f}(u_L) + \mathbf{f}(u_R))$$

$$\begin{aligned}
& - \sum_{\Phi_k \in GNL(\Phi)} \int_0^1 |B|_{FVS}(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \\
& - \sum_{\Phi_k \in LD(\Phi)} \int_0^1 |B|_{FDS}(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.2)
\end{aligned}$$

The numerical flux (3.2) can be equivalently rewritten under the form

$$f(u_L, u_R) = f_{FVS}(u_L, u_R) + \frac{1}{2} \sum_{\Phi_k \in LD(\Phi)} \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds, \quad (3.3a)$$

or

$$f(u_L, u_R) = f_{FDS}(u_L, u_R) - \frac{1}{2} \sum_{\Phi_k \in GNL(\Phi)} \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.3b)$$

Assertion A HUS method is a  $(\Phi; B^\pm)$  scheme in the sense of Definition 1.

The relevance of our definition for a hybridization technique will be assessed in what follows with regards to theoretical as well as practical issues. In particular all the HUS numerical flux-functions, derived below in our new upwinding setting, preserve the Lipschitz-continuity of the underlying FDS and FVS methods. For simplicity in the notation, the endpoints of a given subpath  $\Phi_k(\cdot; u_L, u_R)$  are from now on denoted by

$$\Phi_k(0; u_L, u_R) = u_k; \quad \Phi_k(1; u_L, u_R) = u_{k+1}. \quad (3.4)$$

#### Proof of the Assertion

We have to check that the hybridization procedure introduced above yields upwind methods in the sense of Definition 1. We therefore have to show that the compatibility requirement (2.3) satisfied by both the underlying FDS and FVS methods is preserved when exchanging for a given subpath, i.e., the  $B^\pm$  matrices, according to the nonlinear nature of the associated field.

In this way, let us consider

$$\begin{aligned}
& \int_0^1 (B_{HUS}^+ + B_{HUS}^-)(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = \\
& \sum_{\Phi_k \in GNL(\Phi)} \int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds + \\
& \sum_{\Phi_k \in LD(\Phi)} \int_0^1 (B_{FDS}^+ + B_{FDS}^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.5)
\end{aligned}$$

The first set of integrals related with the GNL subpaths can be rewritten as

$$\int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi(s; u_L, u_R)) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds - \sum_{\Phi_k \in LD(\Phi)} \int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds, \quad (3.6)$$

where in view of the pointwise consistency property (2.42) met by the FVS methods, we have

$$\int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = \mathbf{f}(u_R) - \mathbf{f}(u_L), \quad (3.7)$$

and

$$\sum_{\Phi_k \in LD(\Phi)} \int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \sum_{\Phi_k \in LD(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)). \quad (3.8)$$

Now from the identities (3.6) to (3.8), taking into account our assumption (3.1) for the underlying FDS method clearly enforces the relation (3.5) to reduce to

$$\int_0^1 (B_{HUS}^+ + B_{HUS}^-)(\Phi(s; u_L, u_R), s) \frac{\partial \Phi(s; u_L, u_R)}{\partial s} ds = \mathbf{f}(u_R) - \mathbf{f}(u_L),$$

which is the required result.  $\square$

### Remarks

3.1- As clearly shown by the identity (3.8), the condition (3.1) comes from the use of FVS-type  $B^\pm$  matrices along the GNL subpaths of the FDS-path  $\Phi$ . Such a condition is clearly required for the compatibility condition (2.3) to hold true for the HUS methods. Corollary 1 thus applies and ensures upwind biasing in our framework for hybridization. (3.1) therefore turns out to be a compatibility condition for a given FDS method to provide a suitable candidate for field by field hybridization with FVS schemes.

3.2- In order to stress the relevance of the compatibility condition (3.1), let us assume here that it is not satisfied by the underlying FDS method. We show below that an attempt for hybridization with any given FVS scheme according to (3.2) yields a method that fails in upwind biasing. Let  $f$  denote the resulting numerical flux. We introduce the following formal averaging under the CFL condition (2.4)

$$\tilde{u}_L(u_L, u_R) = u_L - \frac{2h}{h_x} (f(u_L, u_R) - \mathbf{f}(u_L)), \quad (3.9a)$$

$$\tilde{u}_R(u_L, u_R) = u_R - \frac{2h}{h_x} (\mathbf{f}(u_R) - f(u_L, u_R)). \quad (3.9b)$$

Simple algebraic manipulations in (2.5a) and (3.9a) yield:

$$\begin{aligned}
2(f(u_L, u_R) - \mathbf{f}(u_L)) &= 2 \sum_{GNL(\Phi)} \int_0^1 B_{FVS}^-(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \\
&+ 2 \sum_{LD(\Phi)} \int_0^1 B_{FDS}^-(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \\
&+ \sum_{LD(\Phi)} \int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \\
&- \sum_{LD(\Phi)} \int_0^1 (B_{FDS}^+ + B_{FDS}^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.10)
\end{aligned}$$

Thanks to the identity (3.8), we see that the formal averaging (3.9a) does not coincide with the relevant one stated in (2.5a) unless the requirement (3.1) is satisfied; it clearly exhibits unexpected contributions from waves supposed to travel outside of the domain. To go further, assume now that for the states  $u_L$  and  $u_R$  under consideration,  $LD(\Phi) \neq \emptyset$  and that all the states connecting  $u_L$  to  $u_R$  in  $\mathcal{U}$  via  $\Phi(., u_L, u_R)$  satisfy

$$B_{FVS}^-(\Phi(s; u_L, u_R)) = 0, \quad \forall s \in [0, 1], \quad (3.11a)$$

$$B_{FDS}^-(\Phi(s; u_L, u_R), s) = 0, \quad \forall s \in [0, 1]. \quad (3.11b)$$

That is both the FVS and the FDS methods under consideration give for such  $u_L$  and  $u_R$

$$f_{FVS}(u_L, u_R) = f_{FDS}(u_L, u_R) = \mathbf{f}(u_L). \quad (3.12)$$

Then, in view of (3.11), formula (3.10) clearly yields

$$\begin{aligned}
2(f(u_L, u_R) - \mathbf{f}(u_L)) &= \sum_{LD(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)) \\
&- \sum_{LD(\Phi)} \int_0^1 B_{FDS}^+(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \\
&\neq 0.
\end{aligned}$$

Without the condition (3.1),  $f(u_L, u_R)$  does not return the expected flux (3.12). Such a result holds true for any given pair of states  $u_L$  and  $u_R$  that satisfy (3.11) with  $LD(\Phi) \neq \emptyset$ . It is easy to transpose this negative conclusion to (3.9b) to see that conversely when all the signal speeds are nonnegative,  $f$  does not return in general the expected backward flux  $\mathbf{f}(u_R)$ . Such features result in a failure for upwinding. As a consequence, we may expect

the method, which is *not* an HUS method according to our definition, to suffer from a lack of stability. Changing the underlying FVS method is obviously of no help. These developments enlighten the essential role played by the compatibility condition (3.1) for hybridization.

3.3- Since the requirement (2.3) holds true for FDS methods, the condition (3.1) is equivalent to asking for

$$\sum_{\Phi_k \in GNL(\Phi)} \int_0^1 (B_{FDS}^+ + B_{FDS}^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \sum_{\Phi_k \in GNL(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)). \quad (3.13)$$

Indeed, note that we have

$$\sum_{\Phi_k \in GNL(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)) + \sum_{\Phi_k \in LD(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)) = \mathbf{f}(u_R) - \mathbf{f}(u_L).$$

Before investigating for validity the compatibility condition (3.1) for some of the prominent FDS methods, we first state two straightforward consequences of the Definition 3 :

### Corollary 3

*If  $u_L$  and  $u_R$  are connected in the phase space with only linearly degenerate subpaths, that is  $GNL(\Phi) = \emptyset$ ; then any given HUS numerical flux-function (3.2),(3.3b) reduces to*

$$f(u_L, u_R) = f_{FDS}(u_L, u_R). \quad (3.14)$$

*For a single contact discontinuity at rest, we have in particular*

$$f(u_L, u_R) = \mathbf{f}(u_L) = \mathbf{f}(u_R). \quad (3.15)$$

Thus, requirement (P3) is met. Conversely, we also clearly have

### Corollary 4

*If  $u_L$  and  $u_R$  are connected in the phase space with only genuinely nonlinear subpaths, that is  $LD(\Phi) = \emptyset$ ; then any given HUS numerical flux-function (3.2),(3.3a) reduces to*

$$f(u_L, u_R) = f_{FVS}(u_L, u_R). \quad (3.16)$$



In view of the equivalent flux formula (3.3a), the second term on the right hand side may be understood as an antidiffusive procedure to correct a general FVS method with respect to requirement (P3) as stated in Corollary 3. This antidiffusive flux is designed to exactly capture a stationary contact discontinuity and, roughly speaking according to Corollary 4, to vanish for genuinely nonlinear waves in order to restore the ability of the FVS methods for capturing large expansion waves and entropy-satisfying strong shocks.

Note that in the same way, the equivalent flux-formula (3.3b) may be conversely understood with an additional diffusive procedure designed to enhance the robustness of the underlying FDS method in the capturing of nonlinear waves and/or to enforce the selection of physically relevant approximate solutions. Section V devoted to the numerical experiments will in particular assess this important issue.

The above interpretation in terms of an additional antidiffusive flux when considering formula (3.3a) or a diffusive flux, but with respect to (3.3b), must be given a more precise foundation. For such an interpretation to be valid, we should at least expect the following inequalities to be satisfied for  $u_L$  and  $u_R$  in  $\mathcal{U}$  sufficiently close :

$$\int_0^1 \frac{\partial \Phi(t; u_L, u_R)}{\partial t} U_{uu}(\Phi(t; u_L, u_R)) dt \left\{ \sum_{LD(\Phi)} \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \right\} \geq 0,$$

$$\int_0^1 \frac{\partial \Phi(t; u_L, u_R)}{\partial t} U_{uu}(\Phi(t; u_L, u_R)) dt \left\{ \sum_{GNL(\Phi)} \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \right\} \geq 0,$$

where  $U_{uu}$  denotes the Hessian of given entropy  $U$  associated with (1.1). The investigation of these two requirements is however beyond the scope of the present paper and will be addressed in a forthcoming work.

For the current FVS and FDS methods, numerical evidences (see Section V) clearly demonstrate the relevance of the hybridization procedure we propose and strongly support the fact that both the linear and the nonlinear stability requirements (P1),(P2) are met as well.

We now turn to studying the validity of the compatibility requirement (3.1) for some of the prominent FDS methods : namely the Osher-Solomon, the Collela-Glaz and the Roe method. They are derived as a  $(\Phi, B^\pm)$  scheme and will be revisited in Appendix 1 under our formalism. We first state

### Lemma 1

*Both the Osher-Solomon and the Collela-Glaz methods meet the compatibility condition (3.1) for hybridization.*

*Proof of the Lemma 1*

For some of the basic features in use below, we refer the reader to Appendix 1. We consider at first the Osher-Solomon method. By construction, for any given subpath  $\Phi_k \in LD(\Phi)$ , there exists  $\sigma_k \in \mathcal{R}$  such that the endpoints  $u_k = \Phi_k(0; u_L, u_R)$  and  $u_{k+1} = \Phi_k(1; u_L, u_R)$  satisfy the following Rankine-Hugoniot relation

$$\sigma_k(u_{k+1} - u_k) = (f(u_{k+1}) - f(u_k)). \quad (3.17)$$

Here,  $\sigma_k$  is nothing else but the velocity at which the  $k$ th contact discontinuity travels. Now since the method returns (see Appendix 1)

$$\int_0^1 (B^+ + B^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \sigma_k(u_{k+1} - u_k), \quad (3.18)$$

the identity (3.17) clearly ensures the requirement (3.1) to hold true for the Osher-Solomon method.

Again by construction, the Collela-Glaz method does satisfy both the equalities (3.17) and (3.18) for a given  $\sigma_k \in \mathcal{R}$ . The condition (3.1) is thus also valid for this scheme.  $\square$

As a consequence, both the Osher-Solomon and Collela-Glaz methods provide us with relevant candidates for a field by field hybridization. By contrast, we show below that a Roe solver does not necessarily meet the requirement (3.1) for a suitable hybridization. Indeed when considering a Roe type linearization  $B(u_L, u_R)$ , standard arguments yield for a given  $\Phi_k \in LD(\Phi)$  (see again Appendix 1) :

$$\int_0^1 (B^+ + B^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \lambda_k(u_L, u_R)(u_{k+1} - u_k), \quad (3.19)$$

where  $\lambda_k(u_L, u_R)$  denotes the associated eigenvalue coming from the Roe type linearization. Here and by contrast with (3.17),  $(\lambda_k(u_L, u_R); u_k, u_{k+1})$  is solution of the following jump relations

$$\lambda_k(u_L, u_R)(u_{k+1} - u_k) = B(u_L, u_R)(u_{k+1} - u_k). \quad (3.20)$$

Note that we only have *a priori*

$$\sum_{\Phi_k \in LD(\Phi)} \lambda_k(u_L, u_R)(u_{k+1} - u_k) = f(u_R) - f(u_L) - \sum_{\Phi_k \in GNL(\Phi)} \lambda_k(u_L, u_R)(u_{k+1} - u_k)$$

since  $B(u_L, u_R)(u_R - u_L) = f(u_R) - f(u_L)$ . With no further assumption on the Roe type linearization  $B(u_L, u_R)$ , there is no reason for the above expression to reduce to the expected identity (3.1). Indeed, a failure for satisfying (3.1) can be reported for the original

Roe method devoted to the Euler equations. This failure is inherent to the fact that the associated intermediate states do not share in general the same pressure and the same velocity. Thus we have checked that

**Lemma 2** *A general Roe type linearization does not necessarily yield a FDS method that satisfies the compatibility condition (3.1) for hybridization.*

Remarks

3.4- In view of the identity (3.17), we see that both the Osher-Solomon and the Collela-Glaz methods actually obey a stricter compatibility condition than the one required in (3.1). Indeed notice that we have by contrast with (3.1)

$$\int_0^1 (B_{FDS}^+ + B_{FDS}^-)(\Phi_k(s; u_L, u_R), s) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \int_0^1 (B_{FVS}^+ + B_{FVS}^-)(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \mathbf{f}(u_{k+1}) - \mathbf{f}(u_k), \quad (3.21)$$

for any given  $\Phi_k \in LD(\Phi)$ .

3.5- In a recent work devoted to the Euler equations, Mehlman [33] has interpreted the Collela-Glaz method in terms of a Roe solver. Such an interpretation makes nonempty the set of Roe-type linearizations for satisfying the requirement (3.1). Up to our knowledge, the Collela-Glaz method seems to yield the only yet known Roe method that satisfies (3.1) but in a stricter way owing to (3.21). This distinction might leave room for the existence of Roe type linearizations that only obey the weaker form (3.1).

In the sequel, we shall assume that there exist Roe type linearizations with the following additional consistency condition

$$\sum_{\Phi_k \in LD(\Phi)} B(u_L, u_R)(u_{k+1} - u_k) = \sum_{\Phi_k \in LD(\Phi)} (\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)), \quad (3.22)$$

which is nothing but the compatibility condition (3.1) for hybridization. Such an assumption makes the derivation of Roe-type hybridizations a rather formal issue. Nevertheless, such HUS methods exhibit some discrepancies of practical interest with the ones coming from both the Osher-Solomon and the Collela-Glaz schemes. Assumption (3.22) is actually made to provide one with a better insight into the hybridization procedure we propose and to highlight some of its main features.

We now turn to giving a closed-form expression to the numerical flux-function (3.1) resulting from a hybridization of a general FVS method with the FDS methods we have investigated : namely, the Osher-Solomon or the Collela-Glaz methods and the Roe schemes

but subject to (3.22). The statements below indicate the simplicity for formulating a numerical flux for all of these hybridizations of interest.

Concerning the Osher-Solomon and the Collela-Glaz type hybridizations, we have

**Proposition 3**

*Let be given a general FVS method defined by the pair  $(f^+, f^-)$ . For any given  $u_L$  and  $u_R$  in  $\mathcal{U}$ , the numerical flux function associated with the  $(\Phi_{OS}; B_{FVS, OS}^\pm)$  HUS method (respectively the  $(\Phi_{CG}; B_{FDS, CG}^\pm)$  HUS method) reads*

$$f(u_L, u_R) = f^+(u_L) + f^-(u_R) + \sum_{k; \Phi_k \in LD(\Phi)} \epsilon_k (f^{\epsilon_k}(u_{k+1}) - f^{\epsilon_k}(u_k)), \quad (3.23a)$$

with

$$\epsilon_k = -\text{sign}(\sigma_k). \quad (3.23b)$$

Here,  $\sigma_k \equiv \lambda_k(u_k) = \lambda_k(u_{k+1})$  denotes the velocity associated with the  $k$ th linearly degenerate subpath of the Osher-Solomon path (resp. the Collela-Glaz path) (cf. (3.17)).

In a similar way; assuming the existence of relevant Roe methods for hybridization, we state

**Proposition 4**

*Let be given a general FVS method defined by the pair  $(f^+, f^-)$  and a Roe-type linearization  $B(u_L, u_R)$  subject to the additional consistency condition (3.22). For any given  $u_L$  and  $u_R$  in  $\mathcal{U}$ , the numerical flux function associated with the  $(\Phi_{Roe}; B_{FVS, Roe}^\pm)$  HUS method admits the following form*

$$f(u_L, u_R) = f^+(u_L) + f^-(u_R) + \sum_{k; \Phi_k \in LD(\Phi)} \epsilon_k \left\{ (f^{\epsilon_k}(u_{k+1}) - f^{\epsilon_k}(u_k)) + \frac{1}{2} \left( (\lambda_k(u_L, u_R) u_{k+1} - f(u_{k+1})) - (\lambda_k(u_L, u_R) u_k - f(u_k)) \right) \right\}, \quad (3.24a)$$

with

$$\epsilon_k = -\text{sign}(\lambda_k(u_L, u_R)). \quad (3.24b)$$

Here,  $\lambda_k(u_L, u_R)$  denotes the eigenvalue of  $B(u_L, u_R)$  associated with the  $k$ th “linearly degenerate” subpath of the Roe path (cf. (3.20)).

Remarks

3.6- The Osher-Solomon and Collela-Glaz type hybridizations look formally the same. Only, the endpoints  $u_k$  are different.

3.7- By contrast, the Roe-type hybridization admits a somewhat different form of expression. However, this expression is deeply related to the previous ones if we observe

that the Rankine-Hugoniot relations for a  $k$  contact discontinuity ensure when considering (3.17)

$$\sigma_k u_k - \mathbf{f}(u_k) = \sigma_k u_{k+1} - \mathbf{f}(u_{k+1}).$$

Of course, this equality does not hold true for the Roe method which on the other hand satisfies the following *distinct* property of continuity

$$\lambda_k(u_L, u_R)u_k - B(u_L, u_R)u_k = \lambda_k(u_L, u_R)u_{k+1} - B(u_L, u_R)u_{k+1}.$$

already stated in (3.20). These equalities are in general incompatible, except precisely when the  $k$ th contact discontinuity under consideration is the *only* wave separating  $u_L$  and  $u_R$ .

3.8- Note that despite the additional requirement for consistency (3.22) to be met by the Roe methods, we do *not* have in general

$$\sum_{k; \Phi_k \in LD(\Phi)} \epsilon_k \left\{ (\lambda_k(u_L, u_R)u_{k+1} - \mathbf{f}(u_{k+1})) - (\lambda_k(u_L, u_R)u_k - \mathbf{f}(u_k)) \right\} = 0,$$

except precisely when

$$\epsilon_k = \epsilon_l, \quad \forall k, l; \quad \Phi_k, \Phi_l \in LD(\Phi),$$

that is when all the eigenvalues associated with the linearly denenerate subpaths share the same sign, as it is expected in the light of Remark 3.2.

3.9- In view of the above remarks, the discrepancies seen in the formulation of the HUS numerical flux-functions come from the way by which the compatibility condition (3.1) is satisfied by the underlying FDS method. According to the remark 3.4, both the Osher-Solomon and the Collela-Glaz schemes obey the stricter condition (3.21) while the Roe type linearizations under consideration are only expected to satisfy the weaker form (3.1).

Let us stress the fairly simple form resulting from the numerical flux-functions obtained so far. By comparison with the flux formulae (A1.5), (A1.10) and (A1.16) given in Appendix 1, these expressions do not involve the endpoints of any of the inner genuinely nonlinear subpaths, as opposed to all the original approximate Riemann solvers. We also emphasize that sonic points for rarefaction waves play no role in our Osher-type hybridization. As a consequence, the hybrid numerical flux-functions proposed herein are really simpler than their underlying approximate Riemann solvers.

This important practical issue will be further exemplified in the next section devoted to the applications to the Euler equations. Indeed, we shall appeal to the symmetry

property of  $f^\pm$  with respect to the Mach number to directly evaluate  $f^{\epsilon_k}$  and not to deduce  $f^-(u)$  from  $f^+(u)$  through the consistency relation (2.35) as it is commonly done. These observations allow another practical interesting aspect of our new approach for upwinding.

*Proof of Proposition 3 and 4.*

We introduce the following function  $\phi \equiv 2(f_{HVS} - f_{FVS}) : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}^p$ . In view of the flux formula (3.3a), the antidiffusive flux  $\phi$  reads

$$\phi(u_L, u_R) = \sum_{LD(\Phi)} \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.25)$$

Let us first consider the Osher-Solomon type hybridization. We can rewrite  $\phi$  with clear notation as:

$$\phi(u_L, u_R) = \sum_{k; \Phi_k \in LD(\Phi)} \phi_k(u_L, u_R), \quad (3.26a)$$

where

$$\phi_k(u_L, u_R) \equiv \int_0^1 (|B|_{FVS} - |B|_{FDS}) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds. \quad (3.26b)$$

Taking advantage of the Jacobian nature of the  $B_{FVS}^\pm$  matrices, we easily get

$$\begin{aligned} \int_0^1 |B|_{FVS}(\Phi_k(s; u_L, u_R)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = \\ (f^+(u_{k+1}) - f^+(u_k)) - (f^-(u_{k+1}) - f^-(u_k)). \end{aligned} \quad (3.27)$$

On the other hand, thanks to Appendix 1, we have

$$\begin{aligned} \int_0^1 |B|_{FDS}(\Phi_k(s; u_L, u_R, s)) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = |\sigma_k| (u_{k+1} - u_k) \\ = \text{sign}(\sigma_k) \{f(u_{k+1}) - f(u_k)\}. \end{aligned} \quad (3.28)$$

Using the consistency relation (2.35) for the FVS method under consideration, we rewrite

$$f(u_{k+1}) - f(u_k) = \{f^+(u_{k+1}) - f^+(u_k)\} + \{f^-(u_{k+1}) - f^-(u_k)\}, \quad (3.29)$$

so that (3.27), (3.28) and (3.29) yield

$$\phi_k(u_L, u_R) = \begin{cases} -2 (f^-(u_{k+1}) - f^-(u_k)), & \text{if } \sigma_k \geq 0, \\ +2 (f^+(u_{k+1}) - f^+(u_k)), & \text{otherwise,} \end{cases} \quad (3.30)$$

which is the required result.

It is easy to see that this derivation also applies to the Collela-Glaz hybridization along exactly the same lines.

We now come to the Roe-type hybridization. Following the lines of the derivation above, we rewrite  $\phi$  as

$$\phi(u_L, u_R) = \sum_{k; \Phi_k \in LD(\Phi)} \phi_k(u_L, u_R), \quad (3.31a)$$

with

$$\begin{aligned} \phi_k(u_L, u_R) \equiv & \left( \{f^+(u_{k+1}) - f^+(u_k)\} - \{f^-(u_{k+1}) - f^-(u_k)\} \right) \\ & - |\lambda_k(u_L, u_R)|(u_{k+1} - u_k). \end{aligned} \quad (3.31b)$$

Here, we have used the fact that

$$\begin{aligned} \int_0^1 |B(u_L, u_R)|_{FDS} \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds &= |B(u_L, u_R)|_{FDS} (u_{k+1} - u_k) \\ &= |\lambda_k(u_L, u_R)| (u_{k+1} - u_k). \end{aligned} \quad (3.32)$$

To make the function  $\phi_k$  given by (3.31) similar to the one obtained in (3.30), let us write

$$\begin{aligned} |\lambda_k(u_L, u_R)| (u_{k+1} - u_k) &= \text{sign}(\lambda_k) \left( \{ \lambda_k(u_L, u_R) u_{k+1} - f(u_{k+1}) \} \right. \\ &\quad \left. - \{ \lambda_k(u_L, u_R) u_k - f(u_k) \} + \{ f^+(u_{k+1}) - f^+(u_k) \} + \{ f^-(u_{k+1}) - f^-(u_k) \} \right), \end{aligned} \quad (3.33)$$

so that (3.31), (3.32) and (3.33) give the expected result. This concludes the proof.  $\square$

We turn now to study the Lipschitz-continuity of the HUS methods we have considered so far.

### Proposition 5

*Let be given a general FVS method defined by the pair  $(f^+, f^-)$ . For the Osher-Solomon, the Collela-Glaz and the Roe type hybridizations, the associated numerical flux-function respectively given by (3.23) and (3.24) is a locally Lipschitz continuous function  $\mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}^p$ .*

*Proof of Proposition 5.*

Let us again consider the antidiffusive flux  $\phi \equiv 2(f_{HUS} - f_{FVS}) : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}^p$  given by (3.3a). Note that  $f_{HUS}$  and  $\phi$  have the same smoothness.

First, in the case of the Osher-Solomon type hybridization, we recall that  $\phi$  can be written as

$$\phi(u_L, u_R) = \sum_{k; \Phi_k \in LD(\Phi)} \phi_k(u_L, u_R) \quad (3.34a)$$

with

$$\phi_k(u_L, u_R) = \begin{cases} -2(f^-(u_{k+1}) - f^-(u_k)), & \text{if } \sigma_k \geq 0, \\ +2(f^+(u_{k+1}) - f^+(u_k)), & \text{otherwise.} \end{cases} \quad (3.34b)$$

In view of the underlying ODE (1.8) (see also (A1.1)), the endpoints  $u_k(u_L, u_R)$  define smooth mappings  $\mathcal{U} \times \mathcal{U} \rightarrow \mathcal{U}$ . As a consequence, we only have to study the continuity of  $\phi_k$  in the open set  $D_\epsilon^k = \{u_L, u_R \in \mathcal{U}; |\sigma_k| \leq \epsilon\}$  with  $\epsilon$  arbitrary small. By subtracting the two equations in (3.34b), we get

$$2(f^+(u_{k+1}) - f^+(u_k)) + 2(f^-(u_{k+1}) - f^-(u_k)) = 2(\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)).$$

Here we have used the consistency relation (2.35) :  $\mathbf{f}(u_k) = f^+(u_k) + f^-(u_k)$ . For a stationary contact discontinuity, i.e.  $\sigma_k = 0$ , we observe that  $\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k) = 0$ . This proves the required result.

It is easy to see that this conclusion also holds true for the Collela-Glaz type hybridizations. We recall that for this approximate Riemann solver, the endpoints  $u_k(u_L, u_R)$  define locally Lipschitz continuous functions.

We now come to study the smoothness of the Roe-type hybridization. Following the same lines as above, we have to study the smoothness of  $\phi_k(u_L, u_R)$  given, for positive values of  $\lambda_k(u_L, u_R)$ , by

$$\begin{aligned} \phi_k = & -2(f^-(u_{k+1}) - f^-(u_k)) \\ & - \left( (\lambda(u_L, u_R)u_{k+1} - \mathbf{f}(u_{k+1})) - (\lambda_k(u_L, u_R)u_k - \mathbf{f}(u_k)) \right), \end{aligned} \quad (3.35a)$$

and for negative values of  $\lambda_k(u_L, u_R)$ , by

$$\begin{aligned} \phi_k = & 2(f^+(u_{k+1}) - f^+(u_k)) \\ & + \left( (\lambda(u_L, u_R)u_{k+1} - \mathbf{f}(u_{k+1})) - (\lambda_k(u_L, u_R)u_k - \mathbf{f}(u_k)) \right). \end{aligned} \quad (3.35b)$$

We shall assume that the endpoints  $u_k(u_L, u_R)$  associated with the Roe method under consideration, yield locally Lipschitz functions. Therefore, we are again led to study the smoothness of this type of hybridization for  $\lambda_k(u_L, u_R)$  in the neighbourhood of 0. Subtracting again both equalities (3.35a) and (3.35b) yields

$$\begin{aligned} 2(\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)) + 2 \left( (\lambda(u_L, u_R)u_{k+1} - \mathbf{f}(u_{k+1})) \right. \\ \left. - (\lambda_k(u_L, u_R)u_k - \mathbf{f}(u_k)) \right) = 2\lambda(u_L, u_R)(u_{k+1} - u_k). \end{aligned}$$



This ensures the local Lipschitz continuity of the Roe-type hybridization when  $\lambda_k(u_L, u_R) = 0$ . This concludes the proof.  $\square$

### Remarks

3.10- In the proof given above, the enforced consistency condition (3.22) to be met by the Roe-type linearizations plays no role. As a consequence, numerical flux-functions of the form (3.2) but performed with any given Roe method yield consistent  $\mathbf{f}$  and locally Lipschitz continuous functions provided that the Roe method in use meets this last property. However, we again emphasize that without (3.22) and (3.1) the resulting HUS method is not upwind biased.

3.11- In the case of the Roe-type hybridization, it is essential to use the original Roe type linearization to describe the linearly degenerate fields. Exchanging formally the endpoints of the the flux formula (3.23) with the ones coming from the Roe path  $\Phi$ , would yield a slightly simpler numerical flux (cf. expression (3.24)). However, we stress that such a function is *discontinuous*, since in general we have  $\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k) \neq 0$  when  $\lambda_k(u_L, u_R) = 0$  (cf. Remark 3.6), except when  $u_L$  and  $u_R$  are connected in the phase space through a single contact discontinuity at rest. Despite this last property, this simpler procedure (which is *not* a hybrid method) *must not be considered*. We underline that the lack of continuity makes *not valid* the conclusion of the Lax-Wendroff theorem.

These results of Lipschitz-continuity further assess the relevance of our setting proposed for hybrid upstream schemes. In fact, some refinements can be introduced in the light of the previous results. Indeed, we can modify Definition 3 in order to derive simpler Hybrid numerical methods in the case of a path  $\Phi$  made of at least two *distinct* linearly degenerate fields. Let us focus our attention on the Collela-Glaz method for instance, for which the associated approximate Riemann solution is *univalued*. We propose the following procedure:

a- For  $q$ ,  $\Phi_q \in LD(\Phi)$ , find  $q_0$  such that  $\lambda_{q_0} = \min_q |\lambda_q|$  (notice that  $\sigma_q = \lambda_q(u_k) = \lambda_q(u_{k+1})$ ).

b- Define the following numerical  $(\Phi; B^\pm)$  method

For  $k$ ,  $1 \leq k \leq p$  with  $p \neq q_0$ ,  $\Phi_k(s; u_L, u_R)$  is endowed with  $B_{FVS}^\pm$  matrices,

and

For  $k = q_0$ ,  $\Phi_k(s; u_L, u_R)$  is endowed with  $A^\pm$ .

According to this definition, we get

$$\sum_{k \neq q_0} \int_0^1 |B|_{FVS} \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds = (f^+(u_{q_0}) - f^+(u_L)) - (f^-(u_{q_0}) - f^-(u_L))$$

$$+ (f^+(u_R) - f^+(u_{q_0+1})) - (f^-(u_R) - f^-(u_{q_0+1})).$$

On the other hand, we have

$$\int_0^1 |B|_{FDS} \frac{\partial \Phi_{q_0}(s; u_L, u_R)}{\partial s} ds = \text{sign}(\lambda_{q_0})(f(u_{q_0+1}) - f(u_{q_0})).$$

Standard arguments yield the following numerical fluxes

$$f(u_L, u_R) = \begin{cases} f^+(u_L) + f^-(u_R) - (f^-(u_{q_0+1}) - f^-(u_{q_0})), & \text{if } \lambda_{q_0} \geq 0, \\ f^+(u_L) + f^-(u_R) + (f^+(u_{q_0+1}) - f^+(u_{q_0})), & \text{otherwise.} \end{cases}$$

We stress that a stronger compatibility condition than that stated in (3.1) is now required for the above procedure to yield an upwind method : namely the condition (3.21). This condition is precisely satisfied by the Collela-Glaz scheme in use. Also this HUS method is still Lipschitz continuous and satisfies requirement (P3) while being slightly simpler than the one proposed above. On the other hand, this simplification obviously comes at the expense of an enhanced numerical viscosity. Although easy, the extension to the Osher-Solomon path yields a different numerical flux since the associated approximate solution is multi-valued, so that more than one linearly degenerate wave may travel with zero speed.

Moreover, even simpler combinations of FVS and FDS methods can be performed if, in opposition with the Definition 3, the genuinely nonlinear subpaths of  $\Phi$  are not systematically endowed with FVS-type  $B^\pm$  matrices. But it turns out that the resulting simplification always seems to come at the expense of a discontinuous numerical flux function. Once again, we stress that the lack of continuity makes such method an *irrelevant* candidate for practical purposes. In the light of the remarks of this section, the definition we propose appears to be as simple as well as relevant setting for hybrid upstream methods.

#### IV. Application to the Euler Equations

The present section is devoted to applying the general hybridization procedure proposed in Definition 3 to the Euler equations modeling a calorically perfect gas. Detailed flux-formulae are given, coming from a hybridization of a general FVS method with either the Osher-Solomon or the Collela Glaz FDS schemes; these later schemes being revisited in Appendix 1 in terms of our formalism. Let us recall that such FDS methods indeed are relevant for our purpose since they both meet the compatibility requirement (3.1) for hybridization.

Besides the celebrated Van Leer flux vector splitting [48], we shall also focus on the attractive Boltzmann schemes recently introduced by Perthame [38], [39]. Based on the general framework for Boltzmann schemes he proposes, we derive an one-parameter family of such methods, deduced from a previous study by Kaniel [26].

The system of 2D gas dynamic conservation laws projected in an arbitrary coordinate  $x$  becomes

$$\frac{\partial u}{\partial t} + \frac{\partial \mathbf{f}(u)}{\partial x} = 0, \quad (4.1)$$

where the unknown  $u$  and the flux-function  $\mathbf{f}$  read

$$u = \begin{pmatrix} \rho \\ \rho v \\ \rho w \\ \rho E \end{pmatrix}, \quad \mathbf{f}(u) = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ \rho v w \\ (\rho E + p)v \end{pmatrix}. \quad (4.2)$$

Here the thermodynamic closure equation is given by

$$p = (\gamma - 1)\left(\rho E - \frac{1}{2}\rho(v^2 + w^2)\right), \quad \gamma \in [1, 3], \quad \gamma = \text{constant}. \quad (4.3)$$

In the sequel, we give a closed-form expression to various HUS methods based on some of the prominent approximate Riemann solvers. The pair of split fluxes  $(f^+, f^-)$  refers to a general FVS scheme. We shall restrict our attention to FVS methods that satisfy the following properties of symmetry with respect to the Mach number. In particular, both Van Leer's and Perthame's (see Appendix 2) split fluxes  $(f^+, f^-)$  possess this symmetry property. All other state quantities set constant, and  $M = \frac{v}{c}$ ,  $c$  being the speed of sound, we ask, after Van Leer [48], for

$$\begin{aligned} f_\rho^+(M) &= -f_\rho^-(-M), & f_{\rho v}^+(M) &= f_{\rho v}^-(-M), \\ f_{\rho w}^+(M) &= -f_{\rho w}^-(-M), & f_{\rho E}^+(M) &= -f_{\rho E}^-(-M). \end{aligned} \quad (4.4)$$

Note that the exact flux-function (4.2) satisfies such a set of symmetry properties.

## IV.1 HUS Methods Based on Compression/Rarefaction Waves Decomposition

Here, we focus our attention on the Osher-Solomon approximate Riemann solver. For the sake of completeness, some of its basic features are recalled in Appendix 1. The Osher path  $\Phi(\cdot; u_L, u_R)$  (either in the natural or reverse ordering) connects  $u_L$  to  $u_R$  in the physical plane  $(t, x)$  using only rarefaction or compression waves. This path is known to exist in the large as long as there is no cavitation and it is made of respectively two

genuinely nonlinear and two linearly degenerate subpaths for the system (4.1)–(4.3) under consideration. The latter LD subpaths are in fact associated with the contact discontinuity.

Let us denote  $u_L^*$  and  $u_R^*$  the left and right states separated by the contact discontinuity travelling with the velocity  $v^*$ . Let us recall that both the pressure and the velocity  $v$  stay constant across this contact discontinuity. States  $u_L^*$  and  $u_R^*$  are determined by solving the following usual set of equations made of the Riemann invariants of the system considered.

$$\frac{p_L}{\rho_L^\gamma} = \frac{p^*}{\rho_L^{*\gamma}}, \quad w_L = w_L^*, \quad v_L - \epsilon \frac{2}{(\gamma-1)} c_L = v^* - \epsilon \frac{2}{(\gamma-1)} c_L^*, \quad (4.5a)$$

$$\frac{p_R}{\rho_R^\gamma} = \frac{p^*}{\rho_R^{*\gamma}}, \quad w_R = w_R^*, \quad v_R + \epsilon \frac{2}{(\gamma-1)} c_R = v^* + \epsilon \frac{2}{(\gamma-1)} c_R^*, \quad (4.5b)$$

where  $c = \sqrt{\frac{\gamma p}{\rho}}$ . In (4.5a) and (4.5b), we set  $\epsilon = +1$  (respectively  $\epsilon = -1$ ) for a path in the natural order (resp. reverse order). The above system of equations is known to admit a unique and explicit solution (when no cavitation) given by the following algorithm well suited to our purpose:

- 1 ) Compute  $S_L = \rho_L(p_L)^{\frac{-1}{\gamma}}$ ,  $S_R = \rho_R(p_R)^{\frac{-1}{\gamma}}$ , and then  $S_{\frac{1}{2}} = \sqrt{\frac{S_L}{S_R}}$ ,
- 2 ) Compute the sound speed

$$c_L^* = \left[ \frac{(c_R + c_L) + \frac{(\gamma-1)}{2\epsilon}(v_R - v_L)}{1 + S_{\frac{1}{2}}} \right],$$

and deduce  $c_R^* = S_{\frac{1}{2}} c_L^*$ ,

- 3 ) Compute  $v^* = v_L - \epsilon \frac{2}{(\gamma-1)}(c_L - c_L^*)$ ,
- 4 ) Compute  $\kappa = \frac{1}{\gamma} c_L^{*2}$  and  $\pi = (S_L \kappa)^{\frac{1}{(\gamma-1)}}$  and deduce  $\rho_L^* = \pi S_L$ ,  $\rho_R^* = \pi S_R$
- 5 ) Deduce  $p^* = \kappa \rho_L^*$ .

Notice that the above algorithm always returns positive pressure and densities provided that the left and right states meet this property. Moreover besides exponentiations, a fairly small number of divisions are required.

With this solution, we apply the flux formula (3.6a) to the Euler system to get

**Lemma 1** *The HUS numerical flux-function based on a rarefaction/compression waves decomposition reads*

$$f(u_L, u_R) = f^+(u_L) + f^-(u_R) + \begin{cases} -(f^-(u_R^*) - f^-(u_L^*)), & \text{if } v^* \geq 0, \\ +(f^+(u_R^*) - f^+(u_L^*)), & \text{otherwise.} \end{cases} \quad (4.6)$$

Moreover, the switch involved in (4.6) can be automatically accounted for, provided that symmetry conditions (4.4) are satisfied by the underlying FVS method. In that case, the HUS flux reads componentwise

$$f_\rho(u_L, u_R) = f_\rho^+(M_L) + f_\rho^-(M_R) + (f_\rho^+(-|M_R^*|) - f_\rho^+(-|M_L^*|)), \quad (4.7a)$$

$$f_{\rho v}(u_L, u_R) = f_{\rho v}^+(M_L) + f_{\rho v}^-(M_R) - \text{sign}(v^*)(f_{\rho v}^+(-|M_R^*|) - f_{\rho v}^+(-|M_L^*|)), \quad (4.7b)$$

$$f_{\rho w}(u_L, u_R) = f_{\rho w}^+(M_L) + f_{\rho w}^-(M_R) + (f_{\rho w}^+(-|M_R^*|) - f_{\rho w}^+(-|M_L^*|)), \quad (4.7c)$$

$$f_{\rho E}(u_L, u_R) = f_{\rho E}^+(M_L) + f_{\rho E}^-(M_R) + (f_{\rho E}^+(-|M_R^*|) - f_{\rho E}^+(-|M_L^*|)). \quad (4.7d)$$

This significant simplification over the original Osher-Solomon scheme is brought about by the present hybridization procedure. Moreover, no sonic points are involved so only four fluxes need to be evaluated instead of six, and this combination involves no switch.

#### *Proof of Lemma 1*

Starting with the flux formula (3.23a), we rewrite the summation over the two linearly degenerate subpaths  $\Phi_2(u_L^*, u^*)$ ,  $\Phi_3(u^*, u_R^*)$ , where  $u^*$  denotes the intermediate endpoint. Note that these two subpaths are associated with the same eigenvalue  $v^*$ ; we therefore have  $\epsilon_2 = \epsilon_3$ . As a consequence, the intermediate state  $u^*$  plays no role since

$$\begin{aligned} \sum_{\Phi_k, \Phi_k \in LD(\Phi)} \epsilon_k (f^{\epsilon_k}(u_{k+1}) - f^{\epsilon_k}(u_k)) &= \epsilon \left( (f^\epsilon(u_R^*) - f^\epsilon(u^*)) + (f^\epsilon(u^*) - f^\epsilon(u_L^*)) \right), \\ &= \epsilon (f^\epsilon(u_R^*) - f^\epsilon(u_L^*)), \end{aligned} \quad (4.8)$$

which yields the required flux formula (4.6).

Now, thanks to the symmetry properties stated in (4.4), focusing our attention on the first component of the flux (4.7a), we have, setting all other state quantities constant,

if  $v^* \geq 0$ , clearly  $M_L^* = v^*/c_L^* = |M_L^*|$  and the same holds true for  $M_R^*$ ; so that

$$-(f_\rho^-(M_R^*) - f_\rho^-(M_L^*)) = (f_\rho^+(-|M_R^*|) - f_\rho^+(-|M_L^*|)),$$

This identity also holds true for the two last components : namely for  $\rho w$  and  $\rho E$ .

Concerning the momentum equation, note that both the Mach number  $M_L^* = v^*/c_L^*$  and  $M_R^* = v^*/c_R^*$ , share the same sign with  $v^*$ . As a consequence, if  $v^* \geq 0$ , we can write in view of (4.4):

$$-(f_{\rho v}^-(M_R^*) - f_{\rho v}^-(M_L^*)) = -(f_{\rho v}^+(-|M_R^*|) - f_{\rho v}^+(-|M_L^*|))$$

To conclude, when  $v^* \leq 0$ , we notice that  $M_L^* = -|M_L^*|$  and  $M_R^* = -|M_R^*|$ ; therefore the formula (4.7b) is obviously equivalent to (4.6) in that case.

## IV.2 HUS Methods based on Shock/Shock Waves Decomposition

Here we address the hybridization of the Collela-Glaz Riemann solver with a general FVS scheme. As for the Osher path, the present path is made of two genuinely nonlinear subpaths but associated with shock waves and of two linearly degenerate ones related to the contact discontinuity. As mentioned in Remark 3.5, such an hybrid method looks formally the same as those studied in IV.1. Only the endpoints are different. But unlike the rarefaction/compression waves decomposition, no explicit formula is available and an iterative procedure is required to evaluate the corresponding states  $u_L^*$  and  $u_R^*$ . For the sake of completeness, we give below an efficient algorithm for such a purpose.

This algorithm, devised by Gottlieb and Growth [18], solves the problem in considering the following two functions of the velocity  $v$  :  $p_L^*(v)$  and  $p_R^*(v)$  defined as follows:

$$p_L^*(v) = p_L + \gamma p_L \left[ \frac{v - v_L}{c_L} \right] \mathcal{W}_L, \quad (4.9a)$$

where

$$\mathcal{W}_L = \frac{(\gamma + 1)}{4} \left[ \frac{v - v_L}{c_L} \right] - \left( 1 + \left\{ \frac{(\gamma + 1)}{4} \left[ \frac{v - v_L}{c_L} \right] \right\}^2 \right)^{\frac{1}{2}}; \quad (4.9b)$$

and in a symmetric way

$$p_R^*(v) = p_R + \gamma p_R \left[ \frac{v - v_R}{c_R} \right] \mathcal{W}_R, \quad (4.10a)$$

where

$$\mathcal{W}_R = \frac{(\gamma + 1)}{4} \left[ \frac{v - v_R}{c_R} \right] + \left( 1 + \left\{ \frac{(\gamma + 1)}{4} \left[ \frac{v - v_R}{c_R} \right] \right\}^2 \right)^{\frac{1}{2}}. \quad (4.10b)$$

The functions  $p_L^*$  and  $p_R^*$  respectively account for the pressure jump at a given velocity  $v$ , across the left and right shock waves.

These two shock waves can connect  $u_L$  and  $u_R$  via a contact discontinuity provided that there exists a value  $v^*$  of the velocity such that

$$p_L^*(v^*) = p_R^*(v^*) \equiv p^*. \quad (4.11)$$

Note that the pressure  $p^*$  standing for the pressure at the contact discontinuity may be negative. In (4.11),  $v^*$  is nothing but the velocity at which the contact discontinuity travels.

With these notation, the algorithm therefore reads

$$\text{find } v^* \in \mathcal{R}, \text{ such that } p_R^*(v^*) - p_L^*(v^*) = 0. \quad (4.12)$$

A Newton iterative procedure is proved to be really efficient in solving (4.12); the initial guess can be taken as (cf. [18]) :

$$v_0^* = \frac{z\tilde{v}_L + \tilde{v}_R}{1+z} \quad (4.13)$$

where we have set

$$\tilde{v}_L = v_L + \frac{2}{(\gamma-1)}c_L, \quad \tilde{v}_R = u_R - \frac{2}{(\gamma-1)}c_R, \quad z = \frac{c_R}{c_L} \left( \frac{p_L}{p_R} \right)^{\frac{(\gamma-1)}{2\gamma}}. \quad (4.14)$$

Once the velocity  $v^*$  and hence the pressure  $p^*$  are known, densities are given by the following relations

$$\rho_L^* = \rho_L - \frac{(v^* - v_L)^2}{(p^* - p_L)}, \quad \rho_R^* = \rho_R - \frac{(v^* - v_R)^2}{(p^* - p_R)}. \quad (4.15)$$

while the tangential velocities are simply  $w_L^* = w_L$ ,  $w_R^* = w_R$ . The states  $u_L^*$  and  $u_R^*$  are now entirely characterized, the HUS flux-function based on the shock/shock waves decomposition is given by expression (4.6) stated in Lemma 1.

### IV.3 On Some Boltzmann FVS methods

In the hybrid flux formulas derived in the last sections, the pair of split fluxes  $(f^+, f^-)$  refers to a general FVS scheme. Besides the celebrated Van Leer FVS method, the Steger-Warming splitting can be chosen also. The new and attractive Boltzmann-type schemes, considerably improved recently by Perthame in [38], [39], provide one with other pairs of split fluxes  $(f^+, f^-)$ .

Perthame's motivation was to build entropy-satisfying Boltzmann schemes with finite speed of propagation, corresponding to velocity repartition functions with bounded support. Besides the important entropy-satisfying property, the use of velocity repartitions with bounded support brings a dramatic improvement in accuracy over the schemes based on the physical Maxwellian distribution  $e^{-\frac{v^2}{2}}$  (see for instance [10], [43]). Moreover, Perthame has proved under natural conditions concerning the velocity repartitions that both the density and the pressure remains nonnegative [38].

In an independent work, Kaniel [26] has designed such velocity functions with bounded support. These functions yield an one-parameter family of velocity distributions, formally related to the isentropic Euler equations written for a calorically perfect gas with  $\Gamma \in ]1, 3]$ . This family turns out to fall into the framework for Boltzmann schemes proposed by Perthame (see Appendix 2). The splitting is achieved on the equivalent Mach number  $\sqrt{\Gamma p/\rho}$ . The key point is that any given  $\Gamma$ -Boltzmann scheme is naturally made consistent

with the exact flux-function (4.2) for which in general  $\gamma$  is different from  $\Gamma$ . We only consider in this Section the two simplest members corresponding to  $\Gamma = 2$  and  $\Gamma = 3$ . The latter case coincides with a scheme already considered by Perthame [38] and corresponds to the monoatomic one-dimensional gas dynamic. The choice  $\Gamma = 2$  is related to the shallow water equations and to our knowledge, has not been considered by other authors.

These FVS schemes are defined by the pair of split fluxes  $(f^+, f^-)$  we now describe. Let us set

$$c^2 = \frac{\Gamma p}{\rho}; \quad M = \frac{v}{c}, \quad (4.16)$$

where  $c$  and  $M$  respectively stand for a sound speed and a Mach number. Under the CFL like condition

$$\frac{h}{h_x} (|v| + \sqrt{\frac{\Gamma p}{\rho}}) \leq 1, \quad (4.17)$$

the pair  $(f^+, f^-)$  is given by

$$\text{if } M > 1, \quad f^+(u) = \mathbf{f}(u), \quad f^-(u) = 0; \quad (4.18a)$$

$$\text{if } M < 1, \quad f^+(u) = 0, \quad f^-(u) = \mathbf{f}(u); \quad (4.18b)$$

and for  $|M| = \frac{v}{c} \leq 1$ :

$$\begin{array}{c} c \\ |\tilde{M}| \\ f_{\rho}^{\pm} \\ f_{\rho v}^{\pm} \\ f_{\rho w}^{\pm} \\ f_{\rho E}^{\pm} \end{array} \left( \begin{array}{cc} \begin{array}{c} \text{2-Boltzmann} \\ \sqrt{\frac{2p}{\rho}} \\ |M|^3 \\ \pm \frac{1}{6}(2 \pm 3M + |\tilde{M}|)\rho c \\ f_{\rho}^{\pm} v + \frac{1}{6}(3 \pm M(4 - |\tilde{M}|))p \\ f_{\rho}^{\pm} w \\ (\frac{1}{2}v^2 + \lambda \frac{p}{\rho})f_{\rho}^{\pm} \pm \\ \frac{1}{60}(12 \pm 45M + M^2(40 - 7|\tilde{M}|))\rho c \end{array} & \begin{array}{c} \text{3-Boltzmann} \\ \sqrt{\frac{3p}{\rho}} \\ |M|^2 \\ \pm \frac{1}{4}(1 \pm 2M + |\tilde{M}|)\rho c \\ f_{\rho}^{\pm} v + \frac{1}{4}(2 \pm M(3 - |\tilde{M}|))p \\ f_{\rho}^{\pm} w \\ (\frac{1}{2}v^2 + \lambda \frac{p}{\rho})f_{\rho}^{\pm} \pm \\ \frac{3}{16}(1 \pm 4M + M^2(4 - |\tilde{M}|))\rho c \end{array} \end{array} \right), \quad (4.18c)$$

with  $\lambda = \frac{(3-\gamma)}{2(\gamma-1)}$ .

The flux formula given above are expanded for the sake of similarity. Notice however that the 3-Boltzmann scheme can be rewritten under the equivalent but very simple form (see [38])

$$\begin{aligned} f_{\rho}^{\pm} &= \pm \frac{\rho c}{4}(M \pm 1)^2, \quad f_{\rho v}^{\pm} = \pm \frac{\rho c^2}{6}(M \pm 1)^3, \\ f_{\rho w}^{\pm} &= w f_{\rho}^{\pm}, \quad f_{\rho E}^{\pm} = \pm \frac{\rho c^3}{16}(M \pm 1)^4 + \lambda \frac{p}{\rho} f_{\rho}^{\pm} \end{aligned} \quad (4.19)$$



We emphasize after Perthame that the 3-Boltzmann scheme coincides with the Van Leer splitting when  $\gamma = \Gamma = 3$ . Besides the simple expression, these pairs of split fluxes provide us with FVS methods of interest. Let us indeed recall that under the CFL condition (4.17), such FVS methods preserve the positivity of both the density and the pressure [38].

## V. Numerical Experiments

In this section, we investigate the performance of various HUS methods for accuracy and robustness. Intending to illustrate the various properties we have stated, we applied the HUS methods to several test cases ranging from stiff shock-tube problems to 2-D high-speed viscous flows, some of them involving strong shock-wave/boundary layer interaction which leads to flow separation. Such cases provide a convenient benchmark for testing both the shock-capturing capabilities of a given method and its accuracy in the prediction of surface quantities, in particular of the heat transfert rates. Moreover, comparisons of the results are performed with some of the underlying pure FVS and FDS methods in order to enlighten the benefits of the present hybridization technique.

Five distinct sets of problems are addressed to evaluate the relevance of the HUS methods. The first set deals with shock tube problems designed to be a severe test for shock capturing schemes. Results are assessed with the aid of an unusual but enlightening sensitive tool : the entropy balance equation. We point out how this additional equation can be used, numerically speaking, to gain a valuable insight into a general flux splitting method.

The four other sets of problems are to measure the relevance of the HUS methods for Navier-Stokes calculations : namely the one-dimensional conical flow problem proposed by Van Leer, the two dimensional shock wave/boundary layer interaction experimentally investigated by Hakkinen et al., a Mach 25 viscous flow over a blunt body including grid refinement studies and finally two Mach 10 flows past compression ramps with  $15^\circ$  and  $20^\circ$  wedges.

### V.1 Results for the Shock Tube Problems

The problems discussed here are designed for a chemically frozen mixture of calorically perfect gases, diatomic and monoatomic oxygen in the present set of problems; such a modeling makes the specific heat ratio  $\gamma$  depend on the frozen mixture composition and therefore enables distinct but constant  $\gamma$  for the left and right states.

Three distinct HUS methods are tested : the hybridization of the Van Leer FVS scheme with the Osher-Solomon FDS scheme evaluated using the exact Riemann invariants, and the 2- and 3-Boltzmann FVS schemes but with their associated  $\Gamma$ -Riemann invariants.

A grid of uniform 200 points is used for all test cases. The second order of accuracy, when needed, is achieved using the MUSCL approach written in primitive variables. The CFL condition has been set at the constant value 0.5.

#### *Case 1*

*Left state* :  $X_{O_2} = 1, X_O = 0, P = 9.88 \times 10^5, T = 2438, v = 0,$

*Right state* :  $X_{O_2} = 0, X_O = 1, P = 9.93 \times 10^3, T = 2452, v = 0.$

This first test case is designed to show that the use of improper Riemann invariants in the  $\Gamma$ -Boltzmann schemes does not affect their ability to correctly capture strong shock and rarefaction waves. Montagné et al. [34] have reported noticeable errors in the location (and velocity level) of the shock wave in the approximate solutions. To make the problem stiffer, we have deliberately chosen the initial mass fractions in such a way that  $\gamma_L = 1.4$  and  $\gamma_R = 1.6667$  so that a small bump is observed in the velocity profiles, as shown in Fig. 2. The approximate solutions given by the HUS- $\Gamma$ -Boltzmann schemes are quite comparable with the others, thus confirming the relevance of the approach. A closer examination of all results suggests that the HUS with Van Leer and Osher-Solomon hybridization compares the best with the exact solution, specially in the velocity profile.

#### *Case 2*

*Left state* :  $P = 6.50 \times 10^5, T = 2442, v = 0,$

*Right state* :  $P = 1.00 \times 10^3, T = 346, v = 0.$

The mass fractions are calculated under the assumption of chemical equilibrium. The approximate solution calculated by the shock-shock Collela-Glaz approximate Riemann solver at the physical time  $t = 7 \times 10^{-5}s$  is used as the initial condition for this test case. This initial data presents a large nonphysical expansion shock located at  $x = 0.5$ . Our purpose for presenting this problem is to investigate the entropy stability of the HUS methods (hence only looking at the first order accuracy). To this end, we calculated the numerical entropy source term for a given consistent scheme according to the following formulae

$$\tilde{u}_L(u_L, u_R) = u_L - \frac{2h}{h_x}(f(u_L, u_R) - f(u_L)), \quad (5.1a)$$

$$\tilde{u}_R(u_L, u_R) = u_R - \frac{2h}{h_x}(f(u_R) - f(u_L, u_R)), \quad (5.1b)$$

and

$$\begin{aligned} \frac{1}{2}(S(\tilde{u}_L) + S(\tilde{u}_R) - S(u_L) - S(u_R)) \\ + \frac{h}{h_x}(\mathbf{F}(u_R) - \mathbf{F}(u_L)) = \text{Entropy source term} \quad (5.2) \end{aligned}$$

for the convex entropy pair  $(S = \rho \log(\frac{p^{\gamma(x_{O2}, x_{O})}}{p}), \mathbf{F} = S\mathbf{v})$ . The derivation of this formulae is postponed to Appendix 3. As it will be made clear there, formula (5.2) is the entropy equation averaged over a shifted computational cell, centered at each interface ( $CFL \leq 0.5$ ). It accounts for the rate of entropy dissipation/production of a given 3-points scheme for a single jump  $(u_L, u_R)$ . Figure 3 displays the velocity profile at the final physical time  $t = 2.1 \times 10^{-4} \text{ s}$ . All the methods have correctly dealt with the nonphysical shock expansion, except the exact Roe method by Abgrall. The entropy source term is nonpositive, including at the glitches. The plot of the entropy dissipation at  $x = 0.5$  gives a measure of the expected rapid smearing of the inadmissible shock with the time iterations. Both the original and HUS Van Leer methods dissipate entropy at the same rate as the Osher does while the FVS and HUS  $\Gamma$ -Boltzmann schemes yield higher numerical viscosity — neither  $B^+$  nor  $B^-$  admits a vanishing eigenvalue (no glitches are noticeable). We note that any additional viscosity resulting from the energy splitting due to the hybridization formulation only acts in the GNL fields, as clearly shown by the following sets of calculations.

### Case 3

*Left state* :  $P = 3.54 \times 10^4, T = 450, v = 0,$

*Right state* :  $P = 3.49 \times 10^4, T = 2500, v = 0.$

The mass fractions are calculated under the assumption of chemical equilibrium. The exact solution is determined by a slowly moving contact discontinuity at speed  $v^* = 3 \text{ m/s}$ . Our goal is to assess the relevance of our hybridization technique by investigating its numerical diffusion, “inside” a slowly moving contact discontinuity (if exactly at rest, this diffusion vanishes). To this purpose, the entropy source term defined by (5.2) provides a very efficient tool because its magnitude is closely related to the amount of numerical dissipation associated with the method. Figure 4 shows the entropy source term at the location of the contact discontinuity versus time, clearly demonstrating that the numerical dissipation (as measured by the entropy source in (5.2)) of all the HUS methods in the linearly degenerate fields are almost same as that of the Osher scheme and the Roe scheme by Abgrall. By contrast, the dissipation entropy rates for the pure FVS methods are about

30 times higher ! The smearing of the density profiles gives a qualitative measure of the amount of dissipation.

#### *Case 4*

*Left state* :  $P = 5.73 \times 10^2$ ,  $T = 199$ ,  $v = 2200$  ( $M = 7.8$ ),

*Right state* :  $P = 2.23 \times 10^4$ ,  $T = 546$ ,  $v = 0$ .

In this case, we assess the performance of numerical schemes for resolving a shock moving slowly against a large pressure jump (roughly 400 to 1 ratio). Roberts [44] pointed out that a numerical noise could be generated at the shock by some numerical schemes, leaving an oscillation with frequency of several grid points. Figure 5 shows such an oscillation from the Roe and Osher schemes. Both HUS and Van Leer schemes give monotonic resolution of the trailing shock, but a slight bump at the contact discontinuity is seen to be associated with the latter. It is interesting to note that although all schemes capture the trailing shock well with only one to two cells, they perform rather poorly on the leading shock.

#### *Case 5*

*Left state* :  $P = 5.73 \times 10^2$ ,  $T = 199$ ,  $v = 4100$  ( $M = 14.6$ ),

*Right state* :  $P = 5.73 \times 10^2$ ,  $T = 199$ ,  $v = -4000$  ( $M = -14.5$ ).

This case provides a severe test on the capability for resolving shocks resulting from two opposing hypersonic flows. Figure 6 indicates that the Roe, HUS, and Van Leer schemes give very accurate and similar results, except there are slight differences near the initial contact point and at shocks. Both the natural and reverse orderings in the Osher paths for the HUS scheme give comparable results.

## **V.2 Viscous Conical Flow Problem**

Proposed by Van Leer et al. [50], this test case provides an ideal benchmark for testing the accuracy, or conversely the numerical dissipation of a numerical flux-function, for viscous calculations. We solve at the first order of accuracy the one-dimensional conical Navier-Stokes equations for a Mach 7.95 viscous flow with  $\gamma = 1.4$  over a  $10^\circ$  cone, using 64 cells. The Reynolds number is  $Re = 4.2 \times 10^5$  and the Prandtl number is set  $Pr = 0.72$ . Approximate solutions for the temperature, pressure and normal velocity are plotted in Fig. 7 and compared with those obtained by the Roe method. An excellent agreement with this benchmark scheme is achieved for all the HUS method with a slight superiority for the less dissipative Van Leer building block, but only in the capture of the shock wave

(GNL field). By contrast, the pure FVS methods predict considerably broadened boundary layers. These results clearly show the relevance of the hybrid method we propose for viscous calculations. For completeness of evaluating the proposed HUS method based on the Van Leer splitting, we include the convergence test. The present and Roe methods are taking a virtually identical path to convergence, while the Van Leer method is about 50 % faster.

### V.3 Shock Wave/Boundary Layer Interaction Problem

This test provides a comparison of the solutions obtained by the present HUS Van Leer method and Roe scheme with the data measured by Hakkinen et al. The two-dimensional thin-layer laminar Navier-Stokes equations are solved for the flow of Mach 2.0,  $Re = 2.96 \times 10^5$  and  $Pr = 0.72$ . Second-order accurate solutions are carried out on a  $75 \times 65$  grid. As shown in Fig. 8, both HUS and Roe solutions are comparable in accuracy. They give very good agreement with the surface pressure data, while the HUS result is seen to be slightly closer. However, both predict too big a separated region and a long relaxation in the reattaching region. This seems to be a common feature also reported by other authors; but the cause of this discrepancy is yet unknown. Unlike the conical flow case, we discover some interesting features in the convergence history for the HUS. While almost identical in the first four decades of the residual reduction, the HUS method departs from the Roe method by taking a steeper convergence rate, roughly a factor of two.

### V.4 Supersonic/Hypersonic Blunt-Body Flow Problems

The first problem in this category deals with a high-speed laminar flow past a hyperboloid configuration. The curvature radius at the stagnation point is  $R = 0.44 \text{ m}$ , the asymptote angle is given by  $\alpha = 29.3^\circ$ . The flow parameters are  $M_\infty = 25$ ,  $T_\infty = 192.34 \text{ K}$  and  $T_{wall} = 800 \text{ K}$ . The Reynolds number based on the curvature radius at the stagnation point is set to  $Re_R = 7684$ , derived assuming a standard atmosphere at  $H = 77 \text{ km}$ . The Prandtl number is set at  $Pr = 0.72$  and the ratio of specific heats is given by  $\gamma = 1.4$ .

Four grids were used for the grid-convergence study. The most refined grid, made of  $50 \times 112$  nodes, was generated using an exponentially stretched spacing in the streamwise direction while the clustering in the body normal direction is defined using the following ratio parameter  $1 + \epsilon \sin(\pi(j-1)/(nj-1))$ , see Gazaix [15]. From this grid, three subgrids  $50 \times 56$ ,  $50 \times 28$  and  $50 \times 14$  are extracted systematically to study the influence of the refinement in the body normal direction. Figure 9 displays the grid plotted in every other interval in both directions.

Computations with the present HUS method and the underlying Van Leer scheme are compared in terms of surface quantities prediction. Calculations with the Osher method were also carried out on the two finest grids to provide reference predictions. Figure 10 shows the results from the Van Leer method. The pressure coefficient (Fig. 10(a)) is not sensitive to grid resolution. In sharp contrast, the skin friction coefficient (Fig. 10(b)) is heavily affected by increased grid resolution and differs by a significant amount over the entire surface from a grid to another. Mesh resolution has the same drastic effect upon the Stanton number (Fig. 10(c)) : these surface quantities are grossly overpredicted on coarse meshes and significant discrepancies still exist for the two finest grids, indicating that the solution has not yet achieved convergence for the Van Leer method with the finest mesh. These results again confirm the excessive numerical diffusivity of the Van Leer method, already reported by other researchers [14,30,50]. This superfluity in diffusivity finds its roots in the lack of resolution for linear waves (such as contact discontinuity) inherent in the FVS schemes.

The surface quantities with the present HUS method are plotted in Figure 11. Contrary to the Van Leer scheme, no sensitivity with grid resolution is observed for all the surface quantities. The coarsest and finest meshes yield no significant variations in either the skin friction coefficient or the Stanton number. For these two grids, the hardly observed discrepancies are by several orders of magnitude less than the one existing for the FVS quantities computed with the two most refined grids. These results therefore indicate the dramatic improvement in accuracy brought by our hybridization procedure. Besides this significant improvement, the present method exhibits the excellent shock capturing capability from its underlying Van Leer scheme. Indeed, the smoothness of the predicted surface quantities strongly supports that none of these calculations exhibits any tendency towards the development of anomalous solutions, such as the carbuncle phenomena reported for the Roe method [45] or other methods [14].

Turning to the Osher scheme, Figures 12 and 13 depict the predicted skin friction coefficient and the Stanton number compared with the results from the Van Leer and the present HUS methods. The original Osher and HUS schemes predict virtually identical surface quantities for both meshes. This excellent agreement shows that the present hybridization procedure returns the accuracy of the Osher scheme in resolving the boundary layers.

Another problem of interest concerns with the so-called “carbuncle” phenomenon produced by certain schemes such as Roe’s FDS for a supersonic flow over a blunt body, first reported by Peery and Imlay [37]. Figure 14 plots the Mach contours and distribution

of variables along the symmetry line for a Mach 6 flow. The second-order HUS, using Van Leer/Osher splittings, result is free from the “carbuncle” phenomenon and indicates an extremely good shock-capturing capability, with nearly no intermediate shock points. The convergence history displays monotonic decreasing of residual towards machine zero.

## V.5 Hypersonic Compression-Ramp Problems

This set of problems consists of hypersonic laminar flows over  $15^\circ$  and  $20^\circ$  wedges. The flow conditions for the two test cases are the same. The Reynolds number based on the flat plate length  $L$  is  $Re_L = 18000$  and the free stream Mach number  $M_\infty = 10$ . The other parameters are:  $T_\infty = 52\text{ K}$ ,  $T_{wall} = 290\text{ K}$ ,  $Pr = 0.72$ , and  $\gamma = 1.4$ .

These problems have been extensively studied for they were proposed to the second Workshop for Hypersonic Reentry Problems held in Antibes, France in 1991. The Proceedings [40] contains several contributions that provide grid-independent solutions for comparison. A major difference between these two flows is that the  $20^\circ$  ramp angle produces a shock sufficiently strong to cause flow separation of the boundary-layer at the ramp.

These two configurations are calculated on a  $175 \times 50$  grid, see in Fig. 15 the grid for  $15^\circ$  ramp. An analysis of the available results [40] leads us to adopt the following strategy for mesh generation. The space step in the x-direction is set to be constant with  $\Delta x/L = 0.01$ . The vertical spacing is exponentially stretched with a constant parameter less than 1.08, the smallest mesh size being  $\Delta y/L = 0.0002$  at the wall. The corresponding cell Reynolds number  $Re_c \equiv Re_L \Delta y/L$  is 3.6. The downstream end of ramp surface exhibits an expansion corner with a  $-20^\circ$  angle, as proposed by Joulot [25]. This expansion corner is intended to simulate as in the experiment the rapid expansion wave taking place at the end of the compression ramp.

Concerning with first the  $15^\circ$  wedge. The Mach number contours are displayed in Fig. 15b and the surface quantities in Fig. 16, respectively for (a) pressure coefficient, (b) Stanton number, and (c) friction coefficient. Due to the expansion corner, the pressure suddenly drops there and the Stanton number first slightly increases and then drops also. Included also are results obtained using the Osher-Solomon splitting, nearly in complete agreement with the HUS results.

Next turning to the  $20^\circ$  case. In this case, the increased ramp angle causes a flow separation. In Fig. 17, we compare flow quantities at the wall predicted by the HUS method with that by the Osher-Solomon splitting. Both methods again yield virtually

identical profiles. This validates that for the boundary-layer type flow, in which the linear field dominates, the HUS method recovers as intended the accuracy of the Osher-Solomon method.

## VI. Concluding Remarks

In this paper we presented a new upwind method with detailed mathematical construction and analysis. The method is based on a field by field decomposition and endows the *genuinely nonlinear* and *linearly degenerate* waves with effective and compatible pairs of FVS and FDS respectively, thus is termed the *hybrid upwind splitting (HUS)*. We began by defining the compatibility condition for upwind biasing, resulting in a set of path (field)-defined schemes. We showed that the formalism is sufficiently general to include existing approximate Riemann solvers, such as the Roe, the Collela-Glaz, and the Osher-Solomon methods, and the Van Leer and the Steger-Warming flux vector splittings. We gave an interesting interpretation and distinguished the FDS and FVS schemes in terms of the path dependency. Further analysis reveals clearly the direct numerical consequences of the underlying differences of these schemes, specifically their strengthes and weaknesses.

Based upon the formalism, we formulated a suitable hybridization that replaced the weak link of the basic scheme with a compatible path satisfying such desirable properties as: robustness (stability), entropy-satisfying, and low numerical diffusivity. Specifically, the genuinely nonlinear subpaths are endowed with a FVS, and the linearly degenerate one with a FDS. Explicit formulas using the well-known Osher-Solomon, Collela-Glaz, and Roe-type hybridization were given. Unlike the FDS, these expressions do not involve evaluations of endpoints of any of the nonlinear subpaths, nor the sonic point. As a result, the HUS method is simpler (more efficient) than the underlying FDS, but only adding a slight overhead to the FVS in the linearly degenerate subpath, especially as we make use of the symmetry property of the FVS with respect to the Mach number. In addition to the Van Leer splitting, an one-parameter family of the Boltzmann schemes were derived and also included as an alternative of FVS.

To the Euler and Navier-Stokes equations, we applied the HUS methods on several test problems to evaluate their performance against the underlying FDS and FVS methods. These tests convincingly revealed the benefits of the proposed HUS methods, in the respects of stability for solving a strong shock or rarefaction, accuracy for both inviscid and viscous predictions. A formula describing the rate of entropy generation of the numerical schemes was given to provide a valuable information about the associated numerical dissipation, resulting in an unusual but enlightening insight into a general flux splitting method.



The HUS methods in this paper were presented only in the context of explicit schemes. However, issues pertaining to linearization must also be considered as an efficient marching or iterative scheme to achieve converged solution is preferred.

### Acknowledgments

The financial support of ONERA, France and ICOMP while the first author is in residence at ICOMP is acknowledged. The authors would like to thank Drs. Charles E. Feiler of ICOMP, NASA Lewis Research Center, and Yasuhiro Wada of National Aerospace Laboratory, Japan for reading the manuscript. We also thank Drs. Louis A. Povinelli and Robert J. Simoneau of NASA Lewis Research Center for their support of this research.

### Appendix 1

For the sake of completeness, we apply in this appendix the formal transposition of FDS methods in our framework for upwind biasing (proposed in Section II) to some of the prominent approximate Riemann solvers : namely the Osher, the Collela-Glaz and finally the Roe solvers. The following derivations only intend to recall some of their basic features used in Section III. For a complete derivation and furthers results, we refer the reader to the original works.

#### A1.1 The Osher-Solomon Scheme

This upwind method falls by construction into the formalism proposed in Section II. For any given  $u_L$  and  $u_R$  in  $\mathcal{U}$ ,  $\Phi(s; u_L, u_R)$  is defined to be a Lipschitz continuous path made up of  $p$  subpaths, the  $k$ th subpath  $\Phi_k(., u_L, u_R)$  being a solution of the following ODE (see (1.8)) :

$$\frac{d\Phi_k(s; u_L, u_R)}{ds} = r_k(\Phi_k(s; u_L, u_R)). \quad (\text{A1.1})$$

Generally speaking, the Cauchy problem (A1.1) admits an unique solution for  $|u_R - u_L|$  sufficiently small. A given ordering of the  $p$  subpaths yields the initial condition for each of the  $p$  ODE above. Let us underline that the resulting path  $\Phi$  heavily depends on the ordering under consideration. Two orderings have been singled out in [35], namely the natural order  $\Phi = \cup_{k=1}^p \Phi_k$  (see also [42]) and the reverse one  $\Phi = \cup_{k=p}^1 \Phi_k$  put forward by Osher-Solomon for its theoretical advantages (see Section V for a numerical illustration). Actually for the Euler equations and under fairly general thermodynamic assumptions (see [35]), both paths exist in the large as long as there is no cavitation.

Let us denote  $|A| : \mathcal{U} \rightarrow \text{Mat}(\mathcal{R}^p, \mathcal{R}^p)$  the absolute value of the diagonalizable Jacobian matrix  $A$ . Then,  $B^\pm$  are given by

$$B^\pm = \frac{1}{2}(A \pm |A|), \quad (\text{A1.2})$$

which obviously satisfy both assumption (2.2) and the compatibility condition (2.3). The associated numerical flux therefore reads with clear notation :

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \int_{s_k}^{s_{k+1}} |\lambda_k| \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \right), \quad (\text{A1.3})$$

where we have used the following identity deduced from (A1.1)

$$\left( |A| - |\lambda_k| \text{Id}_{\mathcal{R}^p} \right) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} = 0, \quad \text{for } s \in [s_k, s_{k+1}[. \quad (\text{A1.4})$$

Since all the fields are assumed to be either linearly degenerate or genuinely nonlinear, the above vector-valued integration can be explicitly performed and yields

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \text{sign}(\tilde{\lambda}_k) (\mathbf{f}(u_{k+1}) - 2\mathbf{f}(u_{k+\frac{1}{2}}) + \mathbf{f}(u_k)) \right), \quad (\text{A1.5a})$$

where for  $k$ ,  $1 \leq k \leq p$ , we have set

$$\text{sign}(\tilde{\lambda}_k) = \lim_{s \rightarrow s_{k+1}^-} \text{sign}(\lambda_k(\Phi(s; u_L, u_R))), \quad u_k = \Phi(s_k; u_L, u_R). \quad (\text{A1.5b})$$

By convention,  $u_{k+\frac{1}{2}} = u_k$  if  $\lambda_k(u_k)\lambda_k(u_{k+1}) \geq 0$ ; otherwise  $u_{k+\frac{1}{2}}$  is the unique point of  $\Phi_k$  such that  $\lambda_k = 0$ .

## A1.2 The Collela-Glaz scheme

In this approach and by contrast with the previous one, the Riemann problem is solved in the phase space using only (admissible and nonadmissible) shocks and contact waves. The resulting solution is uni-valued (while multi-valued in the Osher-Solomon case, see Dubois [11] and Forestier-Le Floch for a precise definition [13]) and is averaged as in the Godunov method. For  $u_R$  and  $u_L$  close enough, such a solution exists and is made of (at most)  $p$  waves separated by  $(p-1)$  constant states  $u_L \equiv u_0, u_1, \dots, u_p \equiv u_R$  such that there exist  $p$  real numbers  $\sigma_k$  satisfying

$$\sigma_k(u_{k+1} - u_k) = \mathbf{f}(u_{k+1}) - \mathbf{f}(u_k), \quad 1 \leq k \leq p. \quad (\text{A1.6})$$

For any  $u_L$  and  $u_R$  in  $\mathcal{U}$  for which the construction above makes sense, the path  $\Phi(\cdot; u_L, u_R)$  is defined to be  $\cup_{k=1}^p \Phi_k(\cdot; u_L, u_R)$  where for  $k$ ,  $0 \leq k \leq (p-1)$

$$\Phi_k(s; u_L, u_R) \equiv u_k + \frac{(s - s_k)}{(s_{k+1} - s_k)}(u_{k+1} - u_k), \quad \text{for } s \in [s_k, s_{k+1}[ , \quad (\text{A1.7a})$$

with

$$s_0 = 0, \quad s_k = \frac{\sum_{l=1}^k |u_{l+1} - u_l| \mathcal{R}^p}{\sum_{l=1}^p |u_{l+1} - u_l| \mathcal{R}^p}, \quad 1 \leq k \leq p. \quad (\text{A1.7b})$$

This path is locally Lipschitz continuous.

We then define for  $s \in [s_k, s_{k+1}[ , 0 \leq k \leq p$ ,

$$B^\pm(\cdot, s) = \frac{1}{2}(\sigma_k \pm |\sigma_k|)Id_{\mathcal{R}^p}, \quad (\text{A1.8})$$

which clearly satisfy assumption (2.2). Thanks to the Rankine-Hugoniot jump relations (A1.6), we have

$$\sum_{k=1}^p \sigma_k(u_{k+1} - u_k) = \mathbf{f}(u_R) - \mathbf{f}(u_L),$$

so that (2.3) is valid. The numerical flux can be written as

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \int_{s_k}^{s_{k+1}} |\sigma_k| \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \right). \quad (\text{A1.9})$$

An equivalent form is

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \text{sign}(\sigma_k)(\mathbf{f}(u_{k+1}) - \mathbf{f}(u_k)) \right). \quad (\text{A1.10})$$

### Remark

A1.1- From the two constructions above, the Godunov scheme clearly falls into the framework we propose.

### A1.3 The Roe scheme

For two given states  $u_L$  and  $u_R$  in  $\mathcal{U}$ , the Roe method consists in averaging, as in the Godunov method, the exact solution of the following linear problem

$$\frac{\partial u}{\partial t} + \frac{\partial \tilde{\mathbf{f}}(u_L, u_R)(u)}{\partial x} = 0, \quad t > 0, \quad x \in \mathbb{R}, \quad (\text{A1.11a})$$

$$u(0, x) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0. \end{cases} \quad (\text{A1.11b})$$

Here,  $\tilde{\mathbf{f}}(u_L, u_R)(u) = B(u_L, u_R)(u)$ , for all  $u$  in  $\mathcal{U}$ . The matrix  $B(u_L, u_R)$  is called a Roe-type linearization and is required to satisfy the following conditions for any given  $u_L, u_R$  in  $\mathcal{U}$

$$B(u_L, u_R) \text{ is diagonalizable with real eigenvalues,} \quad (\text{A1.12a})$$

$$B(u_L, u_R)(u_R - u_L) = \mathbf{f}(u_R) - \mathbf{f}(u_L), \quad (\text{A1.12b})$$

$$\lim_{u_L \rightarrow u_R} B(u_L, u_R) = A(u_L) = A(u_R). \quad (\text{A1.12c})$$

The exact solution shares the same structure with that considered by Collela-Glaz. The construction we have proposed for that method can therefore be extended in a straightforward way to the Roe method. We just outline this extension. For  $k, 1 \leq k \leq p$ , let  $\lambda_k(u_L, u_R)$  be the  $k$ th eigenvalue of  $B(u_L, u_R)$  and  $r_k(u_L, u_R)$  the associated right eigenvector. Then by considering the following classical decomposition

$$u_R - u_L = \sum_{k=1}^p \nu_k r_k(u_L, u_R), \quad (\text{A1.13})$$

we are led to introduce the following  $(p - 1)$  constant states  $u_L \equiv u_0, u_1, \dots, u_p \equiv u_R$  by setting for  $k, 1 \leq k \leq p$ ,  $u_k = u_{k-1} + \nu_k r_k(u_L, u_R)$ . The path is constructed as in (A1.7), (A1.8),  $B^\pm$  are defined to be

$$B^\pm = \frac{1}{2}(B(u_L, u_R) \pm |B|(u_L, u_R)). \quad (\text{A1.14})$$

The numerical flux associated with the Roe method can be written

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \int_{s_k}^{s_{k+1}} |\lambda_k(u_L, u_R)| \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} ds \right), \quad (\text{A1.15})$$

since we have

$$\left( |B|(u_L, u_R) - |\lambda_k(u_L, u_R)| Id_{\mathcal{R}^p} \right) \frac{\partial \Phi_k(s; u_L, u_R)}{\partial s} = 0, \quad \text{for } s \in [s_k, s_{k+1}]. \quad (\text{A1.16})$$

The Roe numerical flux can be rewritten

$$f(u_L, u_R) = \frac{1}{2} \left( \mathbf{f}(u_L) + \mathbf{f}(u_R) - \sum_{k=1}^p \text{sign}(\lambda_k(u_L, u_R)) (\tilde{\mathbf{f}}(u_L, u_R)(u_{k+1}) - \tilde{\mathbf{f}}(u_L, u_R)(u_k)) \right). \quad (\text{A1.17})$$

Let us stress that the above reconstruction is rather formal. Indeed, the states  $u_k$  associated with a general Roe method do not *a priori* necessarily stay in  $\mathcal{U}$ . In other words,

intermediate states are not necessarily kept within their physical range. We stress that this difficulty is inherent to the Roe method itself.

Although the approximate Riemann solution given by the Roe method is only made of linear waves, for the sake of convenience, a  $k$ th subpath of Roe's path  $\Phi$  will be said to be genuinely nonlinear if the associated eigenvalue  $\lambda_k(u_L, u_R)$  is consistent with the eigenvalue  $\lambda_k(u)$  of  $A(u)$  such that condition (1.5a) is met. Otherwise if (1.5b) is satisfied, then the  $k$ th subpath will be referred to as linearly degenerate.

### Remarks

A1.2- The analogy between the Osher and Roe flux functions, (A1.5) and (A1.17), has been already pointed out by Harten, Lax and Van Leer [22] (see also [36]).

A1.3- In a recent work devoted to multicomponent Euler equations, Mehlman [33] has observed that provided  $u_L$  and  $u_R$  are sufficiently close, then the Collela-Glaz method can be re-expressed equivalently in terms of a Roe-type linearization method.

In view of these remarks, we call attention to two main features of the FDS methods we have revisited in our framework. First, besides their very similar expressions, their formulations are clearly *path dependent*. Then, when  $u_L$  and  $u_R$  are connected in the phase space by a single contact discontinuity with zero speed, all the FDS methods considered above return the exact solution in the following sense

$$\tilde{u}_L(u_L, u_R) = u_L, \quad \tilde{u}_R(u_L, u_R) = u_R \quad (\text{A1.18})$$

since we have (see [4], [35], [45])

$$f(u_L, u_R) = \mathbf{f}(u_L) = \mathbf{f}(u_R). \quad (\text{A1.19})$$

## Appendix 2

In this Appendix, we derive an one-parameter family of FVS methods, we have referred to in Section IV as the  $\Gamma$ -Boltzmann schemes. Here,  $\Gamma$  belongs to  $]1, 3[$ . These schemes come from a set of velocity repartitions parameterized by  $\Gamma$  and designed by Kaniel in [26] to be formally consistent with the one-dimensional Euler equation modeling an isentropic calorically perfect gas, for which the isentropic coefficient is given by  $\Gamma \in ]1, 3[$ . These velocity functions have bounded support and turn out to fall into the general framework due to Perthame [38], [39] for designing entropy- satisfying Boltzmann schemes. Here, we apply this framework to the family of velocity distributions quoted above.

### A2.1 Perthame's kinetic formalism [38]

This section is concerned with the notation and a few basic facts about Perthame methods. Most of the material here is taken from the work of this author, [38] and [37], to which the reader is referred for a comprehensive study and further comments.

Let  $\Theta : \omega \in \mathcal{R} \rightarrow \Theta(\omega) \in \mathcal{R}$  be a nonnegative function satisfying the three consistency conditions

$$\Theta(-\omega) = \Theta(\omega), \quad \int_{\mathcal{R}} \Theta(\omega) d\omega = 1, \quad \int_{\mathcal{R}} \omega^2 \Theta(\omega) d\omega = 1. \quad (\text{A2.1})$$

Moreover, we shall assume that the function  $\Theta$  is boundedly supported, that is, there exists  $\omega_M \in \mathcal{R}^+$  such that

$$\Theta(\omega) = 0, \quad |\omega| \geq \omega_M. \quad (\text{A2.2})$$

In [38], Perthame has shown how to derive from given kinetic entropy pairs such functions  $\Theta$ .

Starting with the given initial data  $u_0(x) = u(0, x)$ , Perthame defines the velocity repartition function associated with this data by

$$\mu_0(x, v) = \frac{\rho(0, x)}{\sqrt{p(0, x)/\rho(0, x)}} \Theta\left(\frac{\omega - v(0, x)}{\sqrt{p(0, x)/\rho(0, x)}}\right), \quad (\text{A2.3a})$$

$$\nu_0(x, v) = \lambda \rho(0, x) \sqrt{\frac{p(0, x)}{\rho(0, x)}} \Theta\left(\frac{\omega - v(0, x)}{\sqrt{p(0, x)/\rho(0, x)}}\right) = \lambda \frac{p(0, x)}{\rho(0, x)} \mu_0(x, v), \quad (\text{A2.3b})$$

with

$$\lambda = \frac{(3 - \gamma)}{2(\gamma - 1)}. \quad (\text{A2.3c})$$

It can be easily checked that the above functions satisfy:

$$\rho(0, x) = \int_{\mathcal{R}} \mu_0(x, \omega) d\omega, \quad \rho w(0, x) = w(0, x) \int_{\mathcal{R}} \mu_0(x, \omega) d\omega, \quad (\text{A2.4a})$$

$$\rho v(0, x) = \int_{\mathcal{R}} \omega \mu_0(x, \omega) d\omega, \quad (\text{A2.4b})$$

$$\rho E(0, x) = \int_{\mathcal{R}} \left\{ \frac{\omega^2}{2} \mu_0(x, \omega) + \nu_0(x, \omega) \right\} d\omega. \quad (\text{A2.4c})$$

Now, consider the following linear transport problems

$$\begin{cases} \partial_t \mu + \omega \partial_x \mu = 0, & t \geq 0, \quad x, \omega \in \mathcal{R}, \\ \partial_t \nu + \omega \partial_x \nu = 0, \\ \mu(0, x, \omega) = \mu_0(x, \omega), \quad \nu(0, x, \omega) = \nu_0(x, \omega), \end{cases} \quad (\text{A2.5})$$

whose exact solutions are explicitly known and given by

$$\mu(t, x, \omega) = \mu_0(x - \omega t, \omega), \quad \nu(t, x, \omega) = \nu_0(x - \omega t, \omega), \quad \forall t \geq 0, \quad x, \omega \in \mathcal{R}. \quad (\text{A2.6})$$

Now let  $h$  be a small time step, then

$$\rho(t, x) = \int_{\mathcal{R}} \mu(t, x, \omega) d\omega, \quad \rho w(t, x) = w(t, x) \int_{\mathcal{R}} \mu(t, x, \omega) d\omega, \quad (\text{A2.7a})$$

$$\rho v(t, x) = \int_{\mathcal{R}} \omega \mu(t, x, \omega) d\omega, \quad (\text{A2.7b})$$

$$\rho E(t, x) = \int_{\mathcal{R}} \left\{ \frac{\omega^2}{2} \mu(t, x, \omega) + \nu(t, x, \omega) \right\} d\omega, \quad (\text{A2.7c})$$

are first-order in  $h$ , approximations of the solution to the Euler equations for  $t \leq h$  (cf. [38]).

Due to the very similarity of the formal construction proposed by Perthame with the collisionless Boltzmann equation, the above velocity repartition functions provide us with a kinetic interpretation for the Euler equations.

A Boltzmann scheme is defined to be a scheme in conservation form equipped with the following numerical flux

$$f(u_L, u_R) = f^+(u_L) + f^-(u_R), \quad \forall (u_L, u_R) \in \mathcal{U} \quad (\text{A2.8})$$

where the split fluxes  $(f^+, f^-)$  are shown to be given for any given piecewise constant data of the form (1.9) by

$$f^+(u) = \int_{\omega \geq 0} \left[ \left( \omega, \omega^2, \frac{\omega^3}{2} \right) \mu_0(x, \omega) + (0, 0, \omega) \nu_0(x, \omega) \right] d\omega, \quad (\text{A2.9a})$$

and

$$f^-(u) = \int_{\omega \leq 0} \left[ \left( \omega, \omega^2, \frac{\omega^3}{2} \right) \mu_0(x, \omega) + (0, 0, \omega) \nu_0(x, \omega) \right] d\omega, \quad (\text{A2.9b})$$

provided that the following CFL like-condition is met :

$$\frac{h}{h_x} (|v(0, x)| + \omega_M \sqrt{\frac{p(0, x)}{\rho(0, x)}}) \leq 1. \quad (\text{A2.10})$$

The condition stated above clearly enforces each velocity distribution of a given computational cell to stay within the adjacent cells.

Notice that summing (A2.9a) with (A2.9b) gives the consistency condition

$$f^+(u) + f^-(u) = \mathbf{f}(u), \quad \forall u \in \mathcal{U}.$$

Boltzmann schemes are shown in [38] to preserve the positivity of the density  $\rho$  and of the pressure  $p$  under the CFL-like condition (A2.10). Moreover, these schemes are entropy satisfying as soon as the generating function  $\Theta$  is deduced from a given kinetic entropy.

## A2.2 Kaniel velocity distribution functions [26]

We now turn to describing a particular set of compactly supported velocity distributions. Proposed by Kaniel [26], these distributions can be actually defined by introducing the following one parameter family of functions  $(\Theta_\Gamma)_{1 < \Gamma \leq 3}$ ,  $\Theta_\Gamma : \omega \in \mathcal{R} \rightarrow \Theta_\Gamma(\omega) \in \mathcal{R}^+$

$$\Theta_\Gamma(\omega) = \frac{1}{\Gamma-1} C_\Gamma |\omega|^{\frac{3-\Gamma}{\Gamma-1}} \mathbf{1}_{[-\sqrt{\Gamma}, +\sqrt{\Gamma}]}(\omega), \quad (\text{A2.11a})$$

where for  $a > 0$ ,  $\mathbf{1}_{[-a, a]}$  denotes the step function

$$\mathbf{1}_{[-a, a]}(\omega) = \begin{cases} 1, & \text{for all } |\omega| \leq a. \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A2.11b})$$

Here, the normalization constant is set to

$$C_\Gamma = \Gamma^{\frac{1}{1-\Gamma}}. \quad (\text{A2.11c})$$

With this choice of  $C_\Gamma$ , it is a simple matter to verify that for any given  $\Gamma \in ]1, 3]$ , we have

$$\begin{aligned} \int_{\mathcal{R}} \Theta_\Gamma(\omega) d\omega &= 2 \int_0^{+\sqrt{\Gamma}} \Theta_\Gamma(\omega) d\omega = 1, \\ \int_{\mathcal{R}} \omega^2 \Theta_\Gamma(\omega) d\omega &= \frac{2\Gamma^{\frac{1}{1-\Gamma}}}{\Gamma-1} \int_0^{+\sqrt{\Gamma}} \omega^{\frac{\Gamma+1}{\Gamma-1}} d\omega = 1, \end{aligned}$$

so that requirements (A2.1) and (A2.2) are satisfied. The one parameter family under consideration therefore yields velocity repartition functions in the sense of the Perthame kinetic formalism.

Here in fact, we had to slightly modify the original functions introduced by Kaniel to make them fall into the framework described above. Instead of (A2.11), the generating function  $\Theta_\Gamma$  is defined in [26] for two given positive real numbers  $\rho$  and  $S$  as follows

$$\Theta_\Gamma[\rho, S](\omega) = \frac{1}{\Gamma-1} C_\Gamma[S] |\omega|^{\frac{3-\Gamma}{\Gamma-1}} \mathbf{1}_{[-\sqrt{\Gamma S \rho^{\Gamma-1}}, +\sqrt{\Gamma S \rho^{\Gamma-1}}]}(\omega), \quad (\text{A2.12a})$$

where the normalization constant  $C_\Gamma[S]$  is chosen to be

$$C_\Gamma[S] = (\Gamma S)^{\frac{1}{1-\Gamma}}. \quad (\text{A2.12b})$$



As a consequence, the resulting  $(1, \omega^2)$  moments of  $\Theta_\Gamma$  are now respectively given by  $\rho$  and  $S\rho^\Gamma$ .

These results allowed Kaniel to define for a given  $\rho$  and  $S$  the following quantity

$$p = S\rho^\Gamma, \quad \text{so that} \quad \left(\frac{\partial p}{\partial \rho}\right)_S = \Gamma S\rho^{\Gamma-1} = \frac{\Gamma p}{\rho} > 0. \quad (\text{A2.13})$$

It is now clear that with respect to the  $\Gamma$ -isentropic Euler equations, the generating function (A2.11) can be understood as a thermodynamic function depending solely on the density  $\rho$  and the entropy  $S$  such that its second moment returns the pressure  $p = S\rho^\Gamma$ . For instance, the choice  $\Gamma = 3$  yields a thermodynamic function consistent with the monoatomic isentropic one-dimensional gas dynamics, while the choice  $\Gamma = 2$  can be formally related to the shallow water equations.

We now derive the explicit flux formulas for the Boltzmann schemes based on the generating functions (A2.11). This is the purpose of the following statement.

**Lemma** *Let be given  $\Gamma \in ]1, 3]$ . Let  $c$  and  $M$  respectively denote an associated sound speed and a Mach number given by*

$$c^2 = \frac{\Gamma p}{\rho}, \quad M = \frac{v}{c}. \quad (\text{A2.14})$$

*Then, under the CFL condition*

$$\frac{h}{h_x} (|v| + \sqrt{\frac{\Gamma p}{\rho}}) \leq 1, \quad (\text{A2.15})$$

*the generating function  $\Theta_\Gamma$  (A2.11) yields a consistent FVS numerical flux defined by the following pair of split fluxes  $(f^+, f^-)$  :*

$$\text{if } M > +1 \quad f^+(u) = \mathbf{f}(u), \quad f^-(u) = 0; \quad (\text{A2.16a})$$

$$\text{if } M < -1 \quad f^+(u) = 0, \quad f^-(u) = \mathbf{f}(u); \quad (\text{A2.16b})$$

*and when  $M \in [-1, 1]$ ,*

$$\begin{aligned} f_\rho^\pm(u) &= \pm \frac{\rho c}{2} \left( \frac{2}{\Gamma+1} \pm M + \frac{\Gamma-1}{\Gamma+1} |M|^{\frac{\Gamma+1}{\Gamma-1}} \right) \\ f_{\rho v}^\pm(u) &= v f_\rho^\pm(u) + \frac{1}{2} \left( 1 \pm \frac{M}{\Gamma+1} (2\Gamma - (\Gamma-1) |M|^{\frac{\Gamma+1}{\Gamma-1}}) \right) p \\ f_{\rho E}^\pm(u) &= \left( \frac{v^2}{2} + \lambda \frac{p}{\rho} \right) f_\rho^\pm(u) \\ &\quad \pm \frac{1}{2} \left( \frac{\Gamma}{3\Gamma-1} \pm \frac{3M}{2} + \frac{M^2}{\Gamma+1} (2\Gamma - \frac{5\Gamma^2 - 8\Gamma + 3}{2(3\Gamma-1)} |M|^{\frac{\Gamma+1}{\Gamma-1}}) \right) pc \end{aligned} \quad (\text{A2.16c})$$

*Proof of the Lemma*

In what follows, we only address the derivation of the forward split flux  $f^+$ . The associated backward split flux is then deduced from the consistency relation (2.35).

Starting with the first component of the forward split flux  $f^+$ , definition (A2.9a) yields

$$f_\rho^+(u) = \frac{C_\Gamma}{\Gamma-1} \frac{\rho}{\sqrt{p/\rho}} \int_{\omega \geq 0} \omega \left| \frac{\omega-v}{\sqrt{p/\rho}} \right|^{\frac{3-\Gamma}{\Gamma-1}} \mathbf{1}_{[-\sqrt{\Gamma}, +\sqrt{\Gamma}]} \left( \frac{\omega-v}{\sqrt{p/\rho}} \right) d\omega. \quad (\text{A2.17})$$

A natural change of variables is given by  $\omega \equiv z\sqrt{\Gamma p/\rho} + v$  and leads us to introduce the following notation

$$c = \sqrt{\frac{\Gamma p}{\rho}}, \quad M = \frac{v}{c} \quad (\text{A2.18})$$

where  $c$  and  $M$  respectively stand for the sound speed and the Mach number given in (A2.14). The above change of variable clearly shows that if  $M < -1$  then  $f_\rho^+(u) = 0$ . Conversely, if  $M > 1$  then  $f_\rho^+(u) = \rho v$ .

Now, focusing our attention on  $M \in [-1, 1]$ , the integral (A2.17) becomes

$$f_\rho^+(u) = \frac{\rho c}{\Gamma-1} \left\{ \int_{-M}^1 z \left| z \right|^{\frac{3-\Gamma}{\Gamma-1}} dz + M \int_{-M}^1 \left| z \right|^{\frac{3-\Gamma}{\Gamma-1}} dz \right\}. \quad (\text{A2.19})$$

The last integral of the right hand-side above is explicitly given by

$$\int_{-M}^1 \left| z \right|^{\frac{3-\Gamma}{\Gamma-1}} dz = \frac{\Gamma-1}{2} (1 + \text{sign}(M) |M|^{\frac{2}{\Gamma-1}}), \quad (\text{A2.20})$$

where by convention,  $\text{sign}(M) = +1$  if  $M > 0$  and  $\text{sign}(M) = -1$  otherwise.

Concerning the first integral in (A2.19), it can be shown that

$$\int_{-M}^1 z \left| z \right|^{\frac{3-\Gamma}{\Gamma-1}} dz = \frac{\Gamma-1}{\Gamma+1} (1 - |M|^{\frac{\Gamma+1}{\Gamma-1}}). \quad (\text{A2.21})$$

Therefore when  $M \in [-1, 1]$ , the mass forward flux reads

$$f_\rho^+(u) = \frac{\rho c}{2} \left( \frac{2}{\Gamma+1} + M + \frac{\Gamma-1}{\Gamma+1} |M|^{\frac{\Gamma+1}{\Gamma-1}} \right)$$

which is the expected result.

Turning to evaluating the second component of the forward split flux, we have

$$f_{\rho v}^+(u) = \frac{C_\Gamma}{\Gamma-1} \frac{\rho}{\sqrt{p/\rho}} \int_{\omega \geq 0} \omega^2 \left| \frac{\omega-v}{\sqrt{p/\rho}} \right|^{\frac{3-\Gamma}{\Gamma-1}} \mathbf{1}_{[-\sqrt{\Gamma}, +\sqrt{\Gamma}]} \left( \frac{\omega-v}{\sqrt{p/\rho}} \right) d\omega. \quad (\text{A2.22})$$

Applying the previous change of variables yields for  $M \in [-1, 1]$

$$f_{\rho v}^+(u) = \frac{\rho c^2}{\Gamma - 1} \left\{ \int_{-M}^1 z^2 |z|^{\frac{3-\Gamma}{\Gamma-1}} dz + 2M \int_{-M}^1 z |z|^{\frac{3-\Gamma}{\Gamma+1}} dz + M^2 \int_{-M}^1 |z|^{\frac{3-\Gamma}{\Gamma-1}} dz \right\}$$

where the first integral is computed as

$$\begin{aligned} \int_{-M}^1 z^2 |z|^{\frac{3-\Gamma}{\Gamma-1}} dz &= \frac{\Gamma - 1}{2\Gamma} (1 + \text{sign}(M) |M|^{\frac{2\Gamma}{\Gamma-1}}) \\ &= \frac{\Gamma - 1}{2\Gamma} (1 + M |M|^{\frac{\Gamma+1}{\Gamma-1}}). \end{aligned} \quad (\text{A2.23})$$

In view of (A2.20) and (A2.21), the momentum numerical flux reduces to the following formula when  $M \in [-1, 1]$  :

$$f_{\rho v}^+(u) = v f_{\rho}^+(u) + \frac{1}{2} \left( 1 + \frac{M}{\Gamma + 1} (2\Gamma - (\Gamma - 1) |M|^{\frac{\Gamma+1}{\Gamma-1}}) \right) p.$$

We now conclude the proof in deriving the last component of the forward split flux. Definition (A2.9a) provides us with

$$\begin{aligned} f_{\rho E}^+(u) &= \frac{C_{\Gamma}}{\Gamma - 1} \frac{\rho}{\sqrt{p/\rho}} \int_{\omega \geq 0} \frac{\omega^2}{2} \left| \frac{\omega - v}{\sqrt{p/\rho}} \right|^{\frac{3-\Gamma}{\Gamma-1}} \mathbf{1}_{[-\sqrt{\Gamma}, +\sqrt{\Gamma}]} \left( \frac{\omega - v}{\sqrt{p/\rho}} \right) d\omega \\ &\quad + \lambda \frac{p}{\rho} f_{\rho}^+(u). \end{aligned} \quad (\text{A2.24})$$

Applying again the same change of variables as above, and after some easy algebraic calculations, we can split the first integral of (A2.24) into

$$\frac{v^2}{2} f_{\rho}^+(u) + \frac{\rho c^3}{2(\Gamma - 1)} \left\{ \int_{-M}^1 z^3 |z|^{\frac{3-\Gamma}{\Gamma-1}} dz + 3M \int_{-M}^1 z^2 |z|^{\frac{3-\Gamma}{\Gamma+1}} dz + 2M^2 \int_{-M}^1 z |z|^{\frac{3-\Gamma}{\Gamma-1}} dz \right\}.$$

Here, the integral to be calculated is explicitly given by

$$\begin{aligned} \int_{-M}^1 z^3 |z|^{\frac{3-\Gamma}{\Gamma-1}} dz &= \frac{\Gamma - 1}{3\Gamma - 1} \left( 1 - |M|^{\frac{3\Gamma-1}{\Gamma-1}} \right) \\ &= \frac{\Gamma - 1}{3\Gamma - 1} \left( 1 - M^2 |M|^{\frac{\Gamma+1}{\Gamma-1}} \right) \end{aligned} \quad (\text{A2.25})$$

Gathering the relations (A2.21), (A2.23) and (A2.25) yields the required expression.  $\square$

### Appendix 3

This appendix is devoted to a brief derivation of the formula (5.2) stated in Section V, accounting for the rate of entropy dissipation/production of a given 3-point scheme. Some comments are given, intending to illustrate its interesting property.

We focus our attention on the system of conservation laws (1.1) that possesses an entropy pair  $(S, \mathbf{F}) : \mathcal{U} \times \mathcal{U} \rightarrow (S(u), \mathbf{F}(u)) \in \mathcal{R} \times \mathcal{R}$ , in the sense of Lax (see [28] for instance) :

$$S \text{ is a convex function of } u, \text{ i.e. } S_{uu} > 0, \quad (\text{A3.1a})$$

$$S \text{ and } \mathbf{F} \text{ satisfy the following compatibility condition } S_u \mathbf{f}_u = \mathbf{F}_u. \quad (\text{A3.1b})$$

In view of the requirement (A3.1b), we see that any given smooth solution  $u$  of (1.1) satisfies the following additional conservation law

$$\frac{\partial S(u)}{\partial t} + \frac{\partial \mathbf{F}(u)}{\partial x} = 0. \quad (\text{A3.2})$$

Weak solutions of (1.1) no longer satisfy this additional conservation law because the jump relation associated with (A3.2) is not compatible, generally speaking, with the Rankine-Hugoniot conditions coming from (1.1), in the sense that it is not satisfied for the same value of the discontinuity velocity. This property yields a suitable criterion for selecting physically relevant weak solutions.

Since weak solutions of (1.1) are not uniquely determined by their initial data (1.2), the physically relevant solutions are selected to be the limit, as the viscosity  $\epsilon$  goes to zero, of solutions  $u_\epsilon$  of the viscous system

$$\frac{\partial u}{\partial t} + \frac{\partial \mathbf{f}(u)}{\partial x} - \epsilon \frac{\partial^2 u}{\partial x^2} = 0, \quad \epsilon > 0. \quad (\text{A3.3})$$

Limit solutions of (A3.3) satisfy in the sense of distributions, the following inequality [20] :

$$\frac{\partial S(u)}{\partial t} + \frac{\partial \mathbf{F}(u)}{\partial x} \leq 0. \quad (\text{A3.4a})$$

That is, if  $u$  is piecewise smooth with discontinuities, then (A3.2) holds pointwise in the smooth regions, while across a discontinuity

$$-\sigma \left\{ S(u_R) - S(u_L) \right\} + \left\{ \mathbf{F}(u_R) - \mathbf{F}(u_L) \right\} \leq 0, \quad (\text{A3.4b})$$

where  $\sigma$  denotes the discontinuity velocity. Relations (A3.4b) are called, after Lax, the entropy conditions. We note that for a shock wave, namely a discontinuity associated with a genuinely nonlinear field, inequality (A3.4b) must be satisfied in a strict manner.

When concerned with the numerical approximation of weak solutions of system (1.1), we must make sure that the method selects only physically relevant solutions, namely the weak solutions that satisfy entropy in the sense of (A3.4). In this way, we ask the method (1.9)-(1.14) to be consistent with the entropy condition in the following sense.

For the method under consideration, there exists a numerical entropy-flux  $F : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{R}$  assumed to be Lipschitz-continuous and consistent with the entropy-flux :

$$F(u, u) = \mathbf{F}(u), \quad \forall u \in \mathcal{U}, \quad (\text{A3.5a})$$

such that the family of approximate solutions  $u^h$  defined in (1.9) satisfies

$$S(u_i^{n+1}) - S(u_i^n) + \frac{h}{h_x} (F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n) \leq 0, \quad i \in \mathcal{Z}, \quad n \in \mathcal{N}, \quad (\text{A3.5b})$$

where we have set

$$F_{i+\frac{1}{2}}^n = F(u_{i+1}^n, u_i^n). \quad (\text{A3.5c})$$

The relevance of requirement (A3.5) is assessed by an extension of the Lax-Wendroff Theorem [20] which states that limits in the sense of bounded,  $L_{loc}^1$  convergence, of approximate solutions given by (1.9)-(1.14) satisfying (A3.5b) are entropy weak solutions of (1.1) and (A3.4).

#### Remark

A3.1- Inequality (A3.5b) can be read as the integral entropy law (A3.4a) applied over the computational cell.

Equipped with these definitions, we see that the entropy condition stated in (A3.5) requires the definition of a numerical entropy-flux, whose derivation actually depends on the numerical flux  $f$  itself. Then according to the above remark, the numerical entropy inequality (A3.5b) is the sum of the entropy rates induced by the averaging (2.5a) and (2.5b) respectively written for the  $(i + 1/2)$  and  $(i - 1/2)$  interfaces. (A3.5b) therefore accounts for two different jump resolutions.

Here, we are merely interested in the entropy rate created in resolving a single jump  $(u_L, u_R)$  in order to avoid a possible misleading mixing of entropy production/dissipation by averaging neighboring (but distinct) contributions. Moreover, we want to circumvent the nontrivial issue of deriving the entropy numerical flux for a specific flux splitting method and instead we intend to arrive at a unique formula as free as possible of the underlying method. These two goals can be achieved in the following way. The derivation below is formal but can be nevertheless rigorously justified on the basis of (A3.5).

We first focus our attention on the half computational cell  $(ih_x, (i+1/2)h_x) \times (nh, (n+1)h)$ . Applying formally the integral conservation law (1.1) over this domain yields under the CFL condition (2.4), the averaging (5.1a) :

$$\tilde{u}_L(u_i^n, u_{i+1}^n) = u_i^n - \frac{2h}{h_x} (f(u_i^n, u_{i+1}^n) - \mathbf{f}(u_i^n)). \quad (\text{A3.6})$$

Again formally, in view of Remark 1, applying now the integral entropy law (A3.4a) over the same domain gives under the CFL condition (2.4)

$$S(\tilde{u}_L(u_i^n, u_{i+1}^n)) - S(u_i^n) + \frac{2h}{h_x} (F(u_i^n, u_{i+1}^n) - F(u_i^n, u_i^n)) \leq 0. \quad (\text{A3.7a})$$

Using the consistency condition (A3.5a), we clearly get :

$$S(\tilde{u}_L(u_i^n, u_{i+1}^n)) - S(u_i^n) + \frac{2h}{h_x} (F(u_i^n, u_{i+1}^n) - \mathbf{F}(u_i^n)) \leq 0. \quad (\text{A3.7b})$$

A similar derivation can be performed also for the rectangle  $((i+1/2)h_x, (i+1)h_x) \times (nh, (n+1)h)$ , providing us successively with

$$\tilde{u}_R(u_i^n, u_{i+1}^n) = u_{i+1}^n - \frac{2h}{h_x} (\mathbf{f}(u_{i+1}^n) - f(u_i^n, u_{i+1}^n)), \quad (\text{A3.8})$$

and

$$S(\tilde{u}_R(u_i^n, u_{i+1}^n)) - S(u_{i+1}^n) + \frac{2h}{h_x} (\mathbf{F}(u_{i+1}^n) - F(u_i^n, u_{i+1}^n)) \leq 0, \quad (\text{A3.9})$$

by using the same arguments as above.

It suffices to sum (A3.7) and (A3.9) to deduce the expected relation (5.2) for the entropy source term

$$\begin{aligned} \frac{1}{2} (S(\tilde{u}_R(u_i^n, u_{i+1}^n)) + S(\tilde{u}_L(u_i^n, u_{i+1}^n)) - S(u_{i+1}^n) - S(u_i^n)) \\ + \frac{h}{h_x} (\mathbf{F}(u_{i+1}^n) - \mathbf{F}(u_i^n)) = \text{Entropy source term}. \end{aligned} \quad (\text{A3.10})$$

Relation (A3.10) clearly accounts for the rate of entropy created in the *shifted* rectangle  $((ih_x, (i+1)h_x) \times (0, h))$  in resolving the single jump  $(u_i^n, u_{i+1}^n)$ . Since such a rate is closely related to the amount of numerical dissipation involved in the resolution of that jump (see [5], [6]), formula (A3.10) provides us with a valuable tool, delivering a clear insight into a general flux splitting method.

## References

- [1] R. Abgrall : *Extension de la Méthode de Roe aux Equations d'Euler Multi-espèces*; La Recherche Aéronautique, vol. 6, pp. 31-43, 1988.
- [2] Y. Brennier : *Averaged Multivalued Solutions for Scalar Conservation Laws*; SIAM J. Numer. Anal., vol. 21, pp. 1013-1037, 1984.
- [3] W.J. Coirier and B. Van Leer : *Numerical Flux Formulas for Euler and Navier-Stokes Equations. Part II : Progress in Flux Vector Splitting*; AIAA paper 91-1566, 1991.
- [4] P. Collela and H.M. Glaz : *Efficient Solution Algorithms for the Riemann Problem for Real Gas*; J. Comp. Phys., vol. 59, pp. 264-289, 1985.
- [5] F. Coquel and P. LeFloch : *Convergence of Finite Difference Schemes for Conservation Laws in Several Space Dimensions : a General Theory*; SIAM J. Numer. Anal., 30, pp. 675-700, 1993.
- [6] F. Coquel and P. LeFloch : *Convergence of Finite Difference Schemes for Conservation Laws in Several Space Dimensions : The Corrected Antidiffusive Flux Approach*; Math. Comp., pp. 169-210, 1991
- [7] F. Coquel and M.-S. Liou : *Hybrid Upwind Splitting Methods*; Proc. of the First European CFD Conference, vol. 1, pp. 9-16, Ch. Hirsh, J. Périaux and E. Onate Editors, Elsevier, 1992.
- [8] F. Coquel and M.-S. Liou : *Field by Field Hybrid Upwind Splitting Methods*; AIAA-93-3302-CP, 1993.
- [9] G. Dal Maso, P. Le Floch and F. Murat : *Definition and Weak Stability of Non-conservative Products*; in P. Le Floch, "Habilitation à diriger des recherches en Mathématiques", Univ. Paris VI, July 1990.
- [10] S. M. Deshpande : *On the Maxwellian Distribution, Symmetric Form, and Entropy Conservation for the Euler Equations*; Tech. Report 2583, NASA Langley, VA, 1986.
- [11] F. Dubois : *Boundary Conditions and the Osher Scheme for the Euler Equations of Gas Dynamics*; Report n° 170, Centre de Mathématiques Appliquées, Ecole Polytechnique, 1987.
- [12] B. Einfeld, C. D. Munz, P. L. Roe and B. Sjoegreen : *On Godunov Type Methods Near Low Densities*; J. Comp. Phys., vol. 92, pp. 273-295, 1991.
- [13] A. Forestier and P. LeFloch : *Multivalued Solutions to Some Nonlinear and Non Strictly Hyperbolic Problems*; Report n° 212, Centre de Mathématiques Appliquées, Ecole Polytechnique, 1990.
- [14] D. Gaitonde and J. S. Shang : *The Performance of Flux Split Algorithms in High-Speed Viscous Flows*, AIAA paper 92-0186, 1992.
- [15] M. Gazaix : *Hypersonic Inviscid and Viscous Flow Computations with a New Optimized Thermodynamic Equilibrium Model*; AIAA paper, 93-0893, Reno, January

1993.

- [16] E. Godlewsky and P. A. Raviart : "Hyperbolic Systems of Conservation Laws", vol. 1 and 2, SMAI (Eds), Ellipses, 1991.
- [17] S. K. Godunov : *A Difference Scheme for Numerical Computation of Discontinuous Solutions of Equations of Fluids Dynamics*; Math. Sbornik, n° 47, pp. 271-306, 1959. (In Russian).
- [18] J. J. Gottlieb and C. P. T. Growth : *Assessment of Riemann Solvers for Unsteady One-dimensional Inviscid Flows of Perfect Gas*; J. Comp. Phys., vol. 78, pp. 437-458, 1988.
- [19] R. J. Hakkinen, I. Grebber, L. Trilling and S. S. Abarnel : *The Interaction of an Oblique Shock Wave with a Laminar Boundary Layer*; NASA Memo 2-18-59W, 1959.
- [20] A. Harten, J. M. Hyman and P. D. Lax : *On Finite Difference Approximations and Entropy Conditions for Shocks*; Comm. Pure and Applied Mathematics, n° 29, pp. 297-322, 1976.
- [21] A. Harten and J. M. Hyman : *Self Adjusting Grid Methods for One-Dimensional Hyperbolic Conservation Laws*; J. Comp. Phys., vol 50, pp. 245-269, 1983.
- [22] A. Harten, P. D. Lax and B. Van Leer : *On Upstream Differencing and Godunov-type Schemes for Hyperbolic Conservation Laws*; SIAM Review, vol. 25, pp. 35-61, 1983.
- [23] Ch. Hirsch : *Numerical Computation of Internal and External Flows*; Volumes 1 and 2, John Wiley Editor.
- [24] P. A. Jacobs : *Single-Block Navier-Stokes Integrator*; ICASE Interim Report 18, July 1991.
- [25] A. Joulot : in Proc. of the Workshop on hypersonic flows for reentry problems, Part II, vol. 3, pp. 1-20, April 15-19, 1991, Antibes, France.
- [26] S. Kaniel : *A Kinetic Model for the Compressible Flow Equations*; Indiana Univ. Math. J., vol. 37, n° 3, pp. 537-563, 1988.
- [27] B. Larrouturou : Proc. of the first European CFD conference, vol. of invited lectures, pp. 117-126, Ch. Hirsh, J. Périaux and E. Onate Editors, Elsevier, 1992.
- [28] P. D. Lax : *Shock waves and Entropy*; In E. H. Zarantonello (Ed.), Contribution to Nonlinear Functional Analysis, pp. 603-634, New York : Academic Press, 1971.
- [29] A. Lerat : *Propriétés d'Homogénéité et Décomposition des Flux en Dynamique des Gaz*; Journal de Mécanique Théorique et Appliquée, n° 2, pp. 185-213, 1983.
- [30] M.-S. Liou and C. Steffen : *A New Flux Splitting Scheme*; J. Comp. Phys., vol 107, pp. 23-39, 1993.
- [31] M.-S. Liou and C. Steffen : *HOPE : High Order Polynomial Expansions Flux Vector Splitting Methods*; NASA TM-104452, 1991.



- [32] M.-S. Liou : *A Continuing Search for a Near-Perfect Numerical Flux Scheme, Part I: AUSM<sup>+</sup>*; NASA TM-106524, 1994.
- [33] G. Mehlman : *An Approximate Riemann Solver for Fluid Systems Based on a Shock Curve Decomposition*; Proceedings of the 3<sup>rd</sup> International conference on Hyperbolic Problems, Engquist Gustafsson eds., Chartwell Bratt, pp. 727-741, 1991.
- [34] J. L. Montagné, H. C. Yee, and M. Vinokur : *Comparative Study of High Resolution Shock Capturing Schemes for Real Gases*; NASA TM-86839, 1987.
- [35] S. Osher and F. Solomon : *Upwind Difference Schemes for Hyperbolic Systemes of Conservation Laws*; Math. of Comp., vol. 38, n° 158, pp. 339-374, 1982.
- [36] S. Osher : *Riemann Solvers, the Entropy Condition, and Difference Approximations*; SIAM J. Numer. Anal., vol. 21, pp 217- , 1984.
- [37] K. M. Peery and S. T. Imlay : *Blunt-Body Flow Simulations*; AIAA paper 88-2904, AIAA/SAE/ASME/ASEE 24th Joint Propulsion Conference, 1988.
- [38] B. Perthame : *Boltzmann Type Schemes for Gas Dynamics and the Entropy Property*; SIAM J. Numer. Anal., vol 27, n° 6, pp. 1405-1421, 1990.
- [39] B. Perthame : *Second-Order Boltzmann Schemes for Compressible Euler Equations in One and Two Space Dimensions*; SIAM J. Numer. Anal., vol 29, n° 1, pp. 1-19, 1992.
- [40] Proceedings of the *Workshop on hypersonic flows for reentry problems*, Part II, April 15-19, 1991, Antibes, France.
- [41] J. J. Quirk : *A Contribution to the Great Riemann Solver Debate*; ICASE Report 92-64, 1992.
- [42] M. M. Rai and S. Chakravarthy : *An Implicit Form for the Osher Upwind Scheme*; AIAA Paper-84-0088, 1984.
- [43] R. D. Rietz : *One-Dimensional Compressible Gas Dynamics Calculations Using the Boltzmann Equation*; J. Comp. Phys., n° 42, pp. 108-123, 1981.
- [44] T. W. Roberts : *The behavior of Flux Difference Splitting Schemes near Slowly Moving Shock Waves*; J. Comput. Phys., vol. 90, pp. 141-160, 1990.
- [45] P. L. Roe : *Approximate Riemann Solvers, Parameter Vectors and Difference Schemes*; J. Comp. Phys., vol. 43, pp. 357-372, 1981.
- [46] P. L. Roe : *Upwind Schemes Using Various Formulations of the Euler Equations*; In F. Angrand et al. (eds), Numerical Methods for the Euler Equations of Fluid Dynamics, Philadelphia, SIAM Publications, 1983.
- [47] J. L. Steger and R. F. Warming : *Flux Vector Splitting of the Inviscid Gas-Dynamic Equations with Applications to Finite Difference Methods*; J. Comp. Phys., n° 40, pp263-293, 1981.

- [48] B. Van Leer : *Flux Vector Splitting for the Euler Equations*; Proc. 8th International Conference on Numerical Methods in Fluid Dynamics, Berlin : Springer Verlag, 1982.
- [49] B. Van Leer : *On the Relation Between the Upwind Differencing Schemes of Godunov, Engquist-Osher and Roe*; SIAM Journal of Sci. Stat. Computing, n° 5, pp. 1-20, 1984.
- [50] B. Van Leer, J. L. Thomas, P. L. Roe and R. W. Newsome : *A Comparison of Numerical Flux Formulas for the Euler and Navier-Stokes Equations*; AIAA paper-87-1104-CP, 1987.
- [51] Y. Wada and M.-S. Liou : *A Flux Splitting Scheme with High-Resolution and Robustness for Discontinuities*; AIAA paper 94-0083, 1994; also NASA TM 106452, 1994.

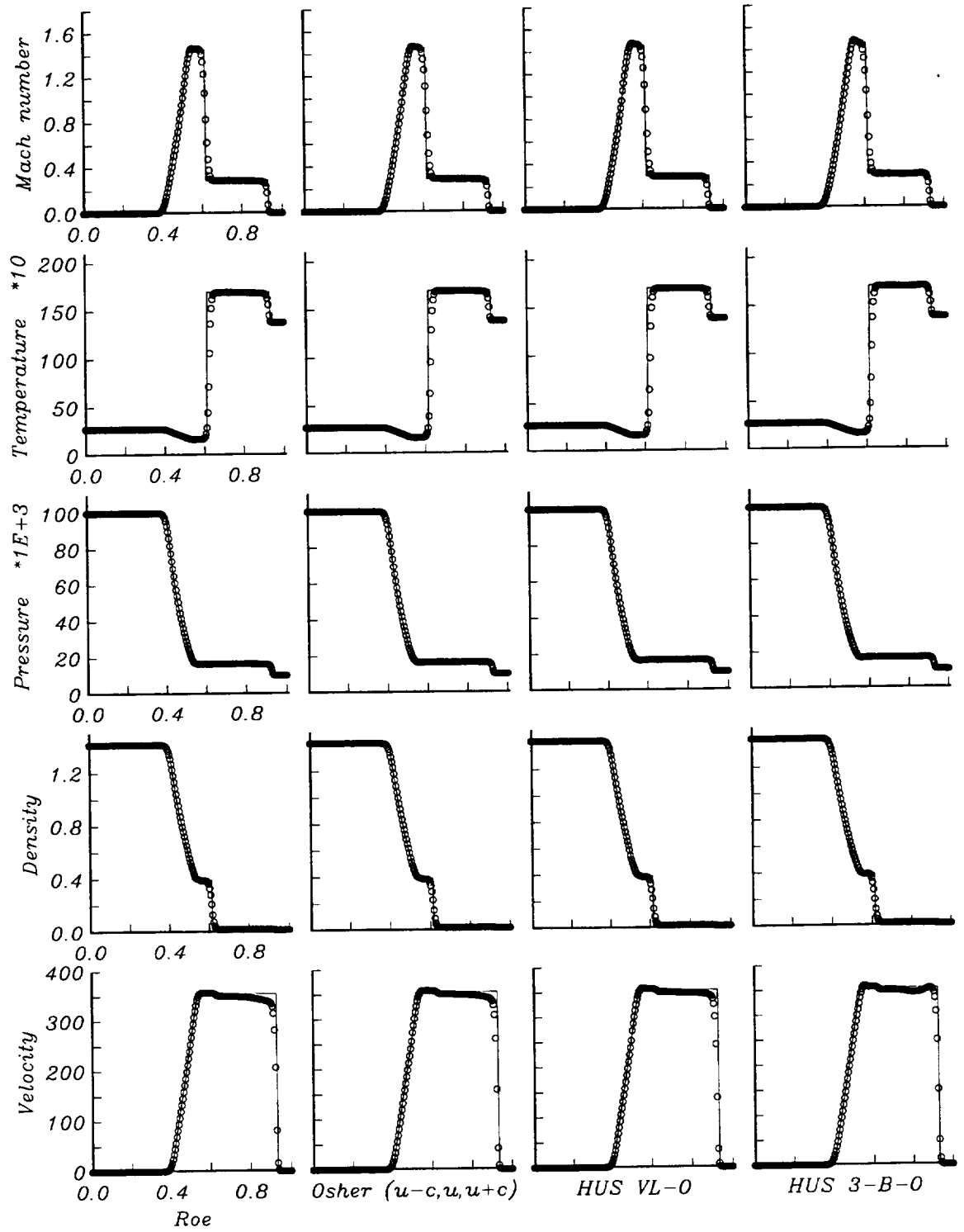


Figure 2. Shock tube problem, Case 1: (a) Comparison of the results from the methods of Roe, Osher (natural ordering), HUS (Van Leer-Osher hybridization), HUS (Boltzmann-Osher hybridization with  $\Gamma = 3$ ), and the exact Riemann solver (solid line).

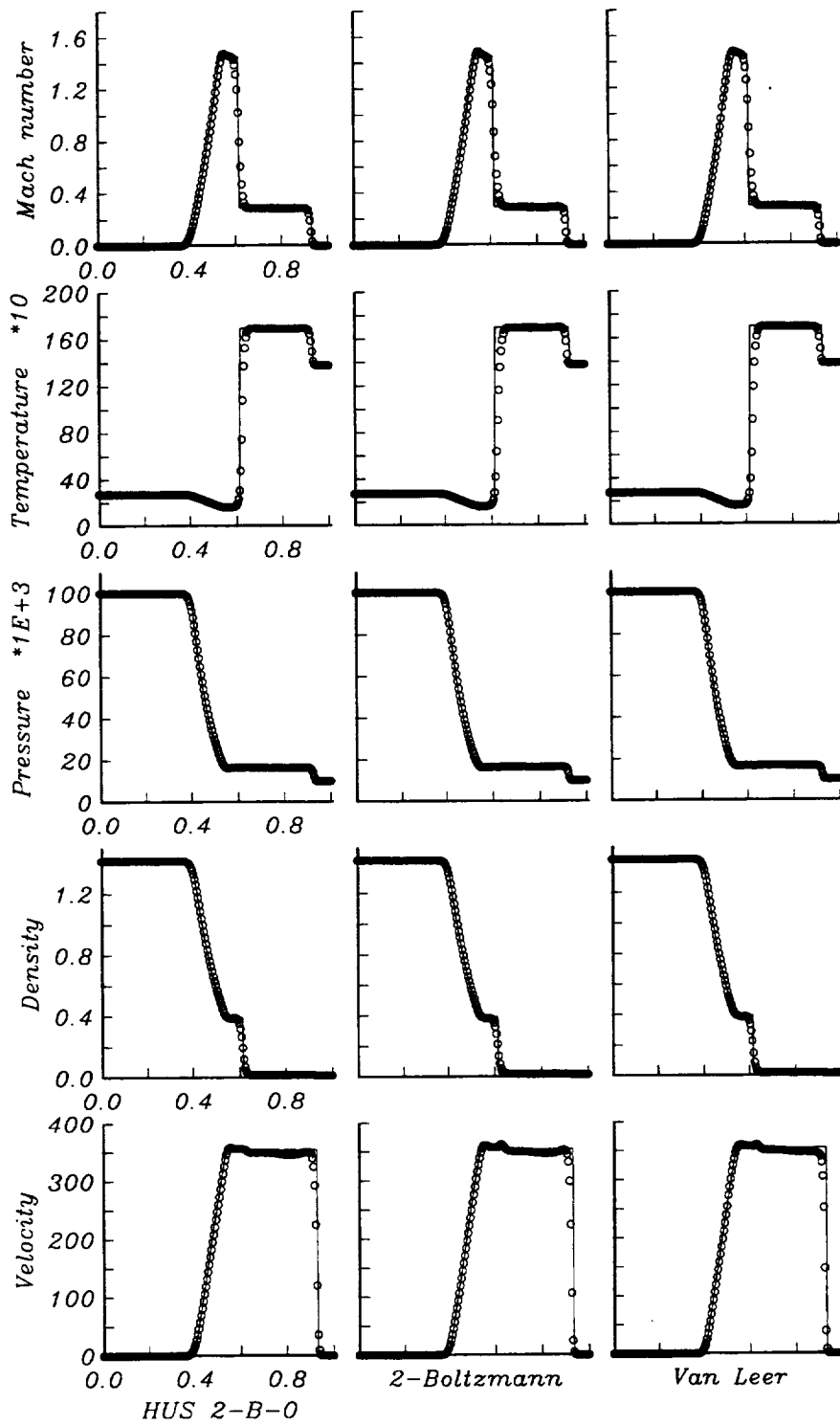


Figure 2. (continued) (b) Comparison of results from the methods of HUS (Boltzmann-Osher hybridization with  $\Gamma = 2$ ), Boltzmann ( $\Gamma = 2$ ), Van Leer and the exact Riemann solver (solid line).

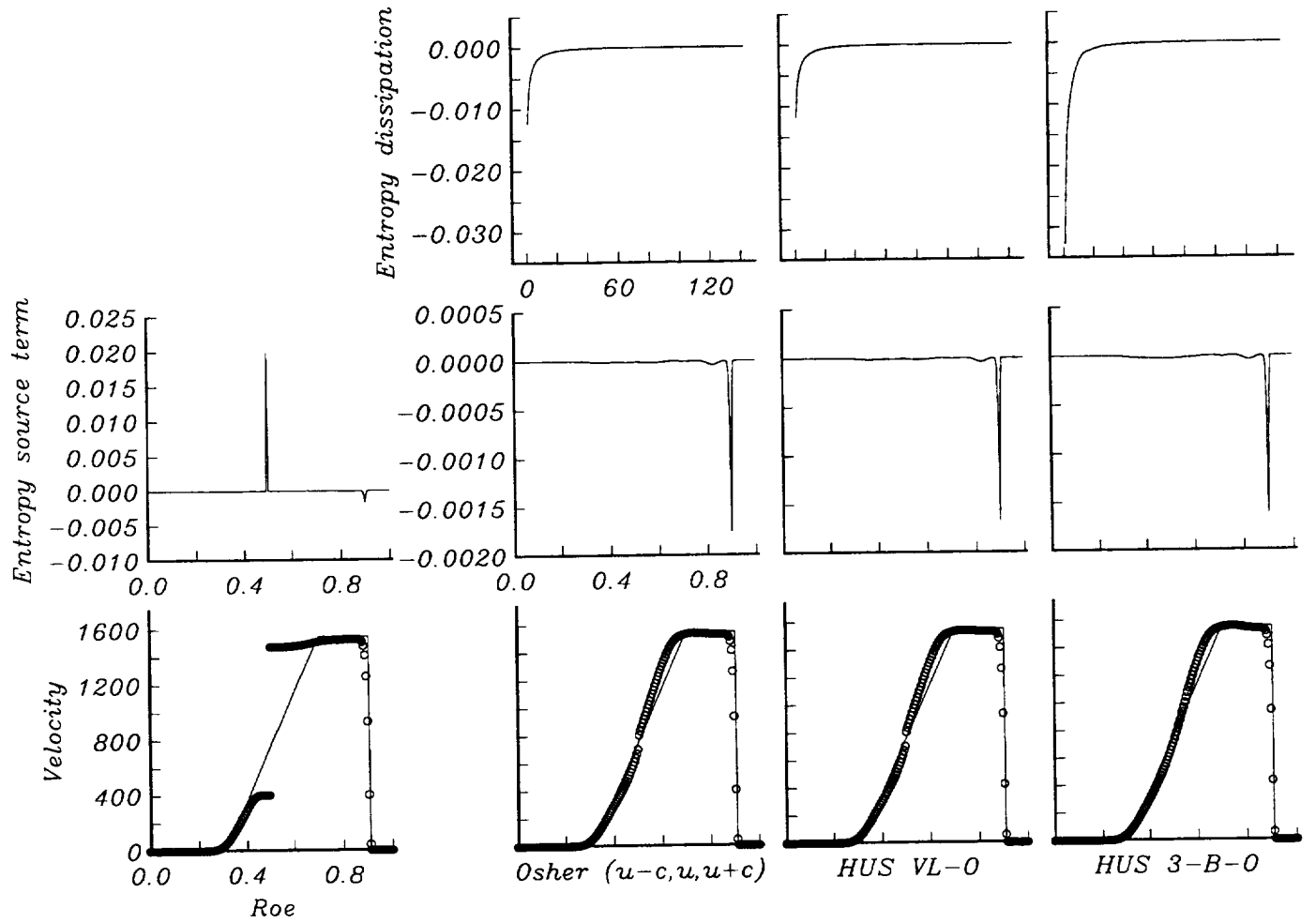
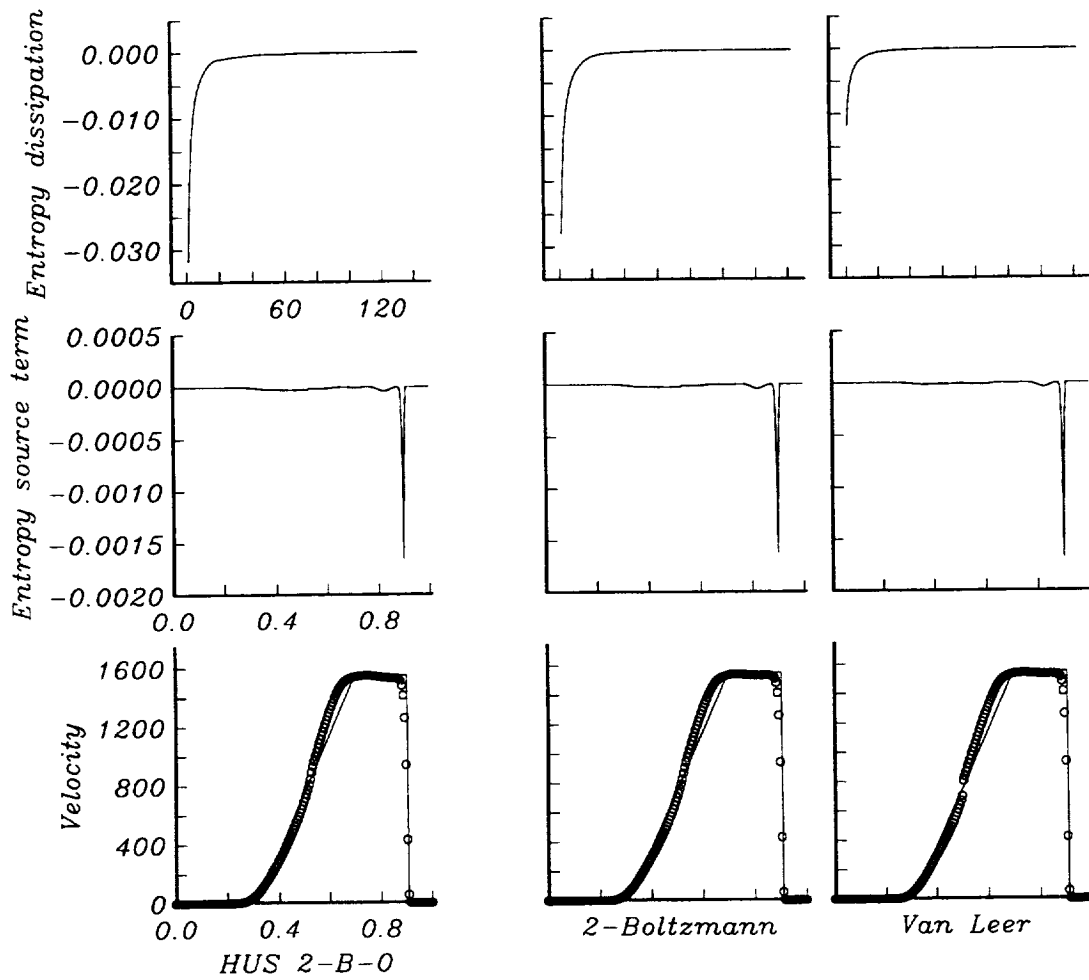


Figure 3. Shock tube problem, Case 2: (a) Comparison of results from the methods of Roe, Osher (natural ordering), HUS (Van Leer-Osher hybridization), HUS (Boltzmann-Osher hybridization with  $\Gamma = 3$ ) and the exact Riemann solver (solid line). The entropy source term is defined in (5.2) and the entropy dissipation is the time history of this term measured at  $x = 0.5$ .



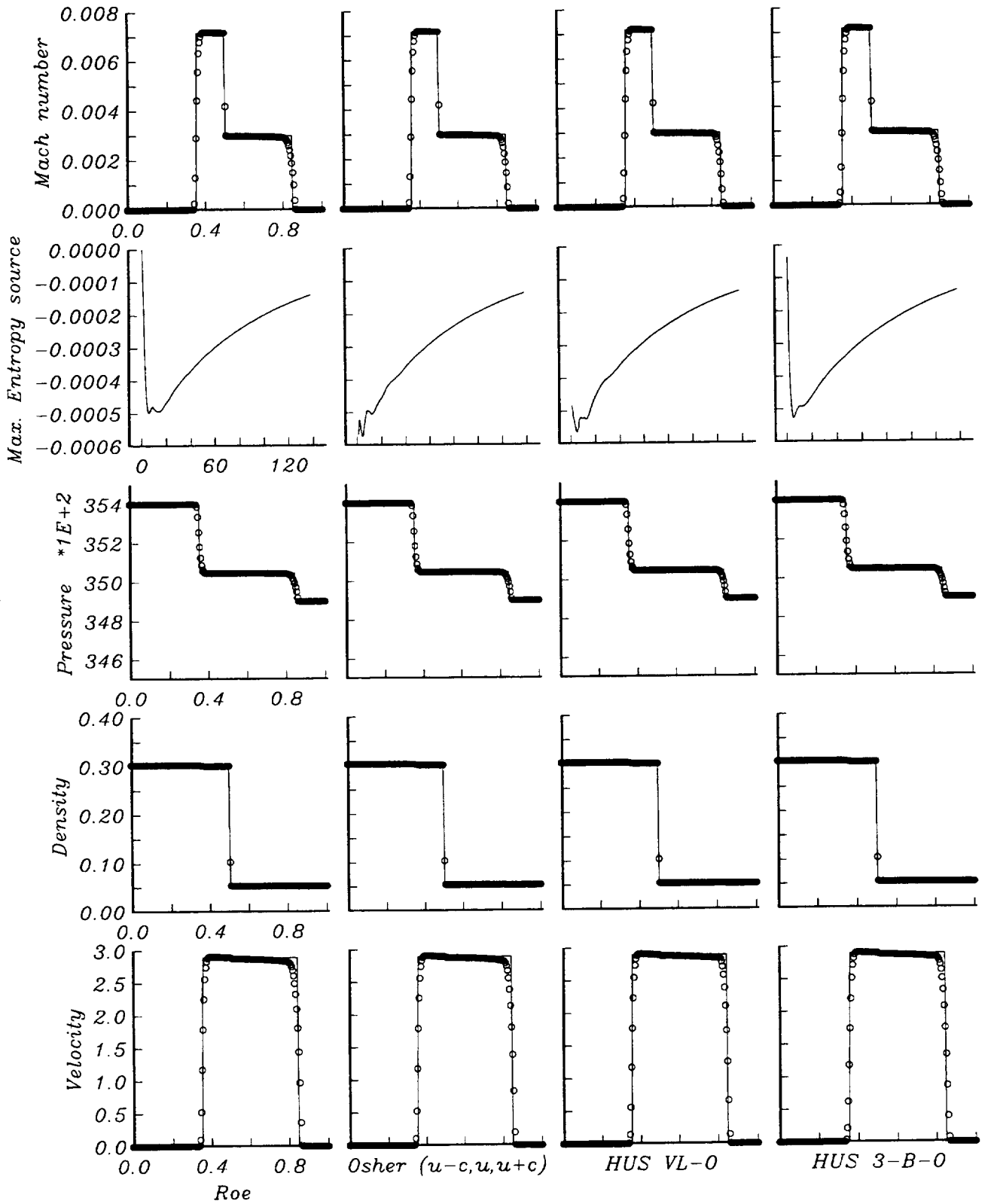


Figure 4. Shock tube problem, Case 3: (a) Comparison of results from the methods of Roe, Osher (natural ordering), HUS (Van Leer-Osher hybridization), HUS (Boltzmann-Osher hybridization with  $\Gamma = 3$ ), and the exact Riemann solver (solid line).

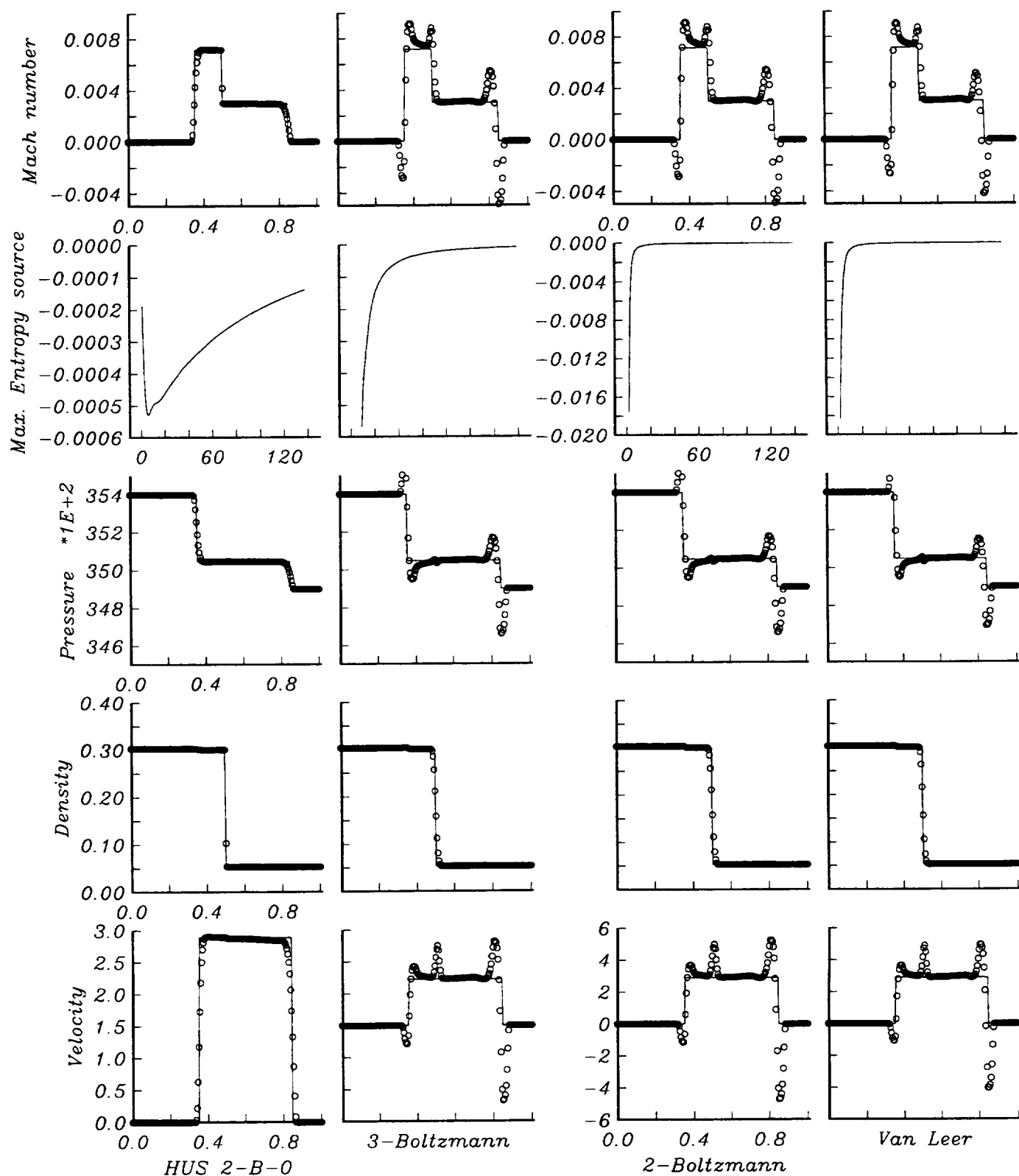


Figure 4. (continued) (b) Comparison of results from the methods of HUS (Boltzmann-Osher hybridization with  $\Gamma = 2$ ), Boltzmann ( $\Gamma = 3$ ), Boltzmann ( $\Gamma = 2$ ), Van Leer and the exact Riemann solver (solid line).



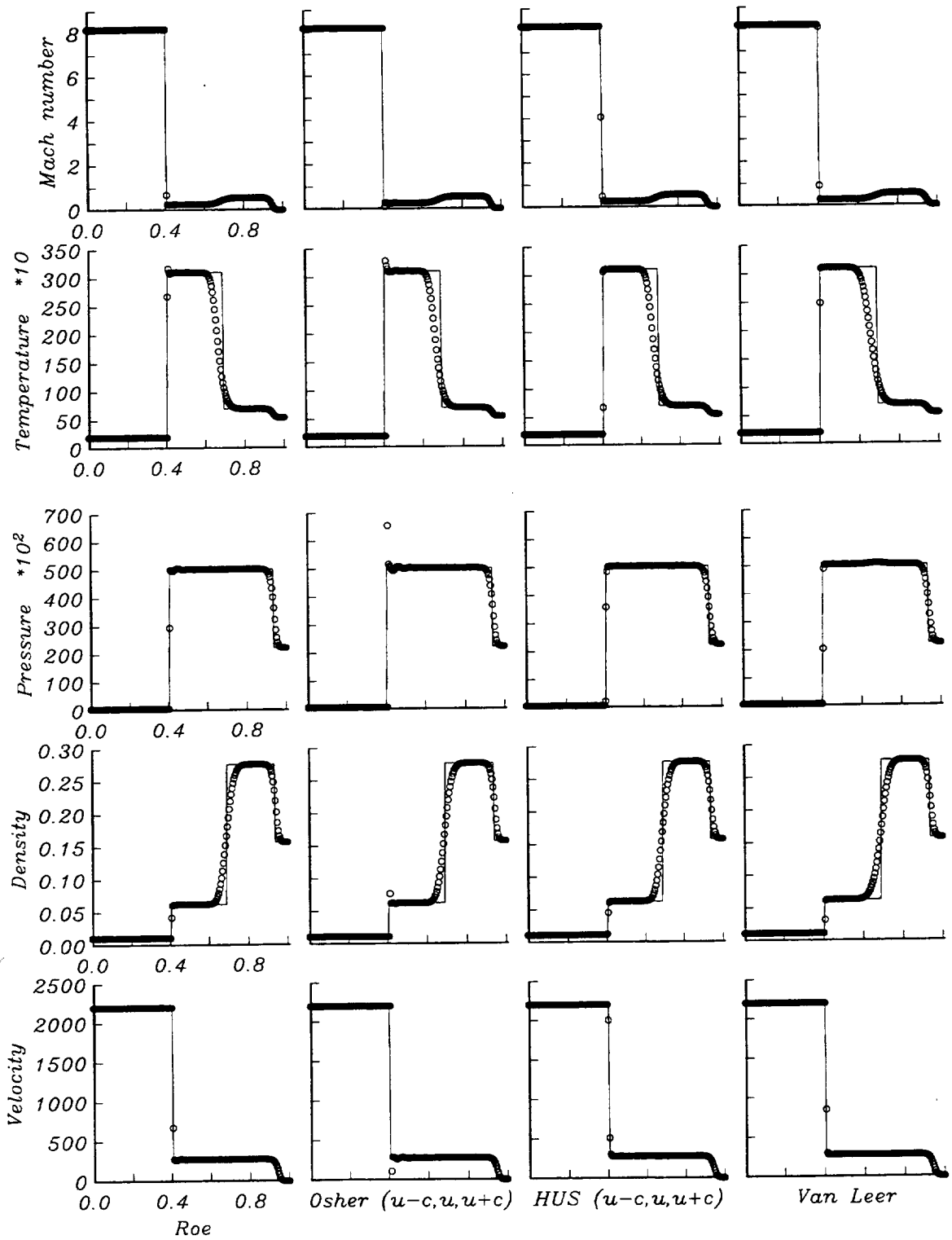


Figure 5. Shock tube problem, Case 4: (a) Comparison of results from the methods of Roe, Osher (natural ordering), HUS (Van Leer-Osher hybridization), Van Leer, and the exact Riemann solver (solid line).

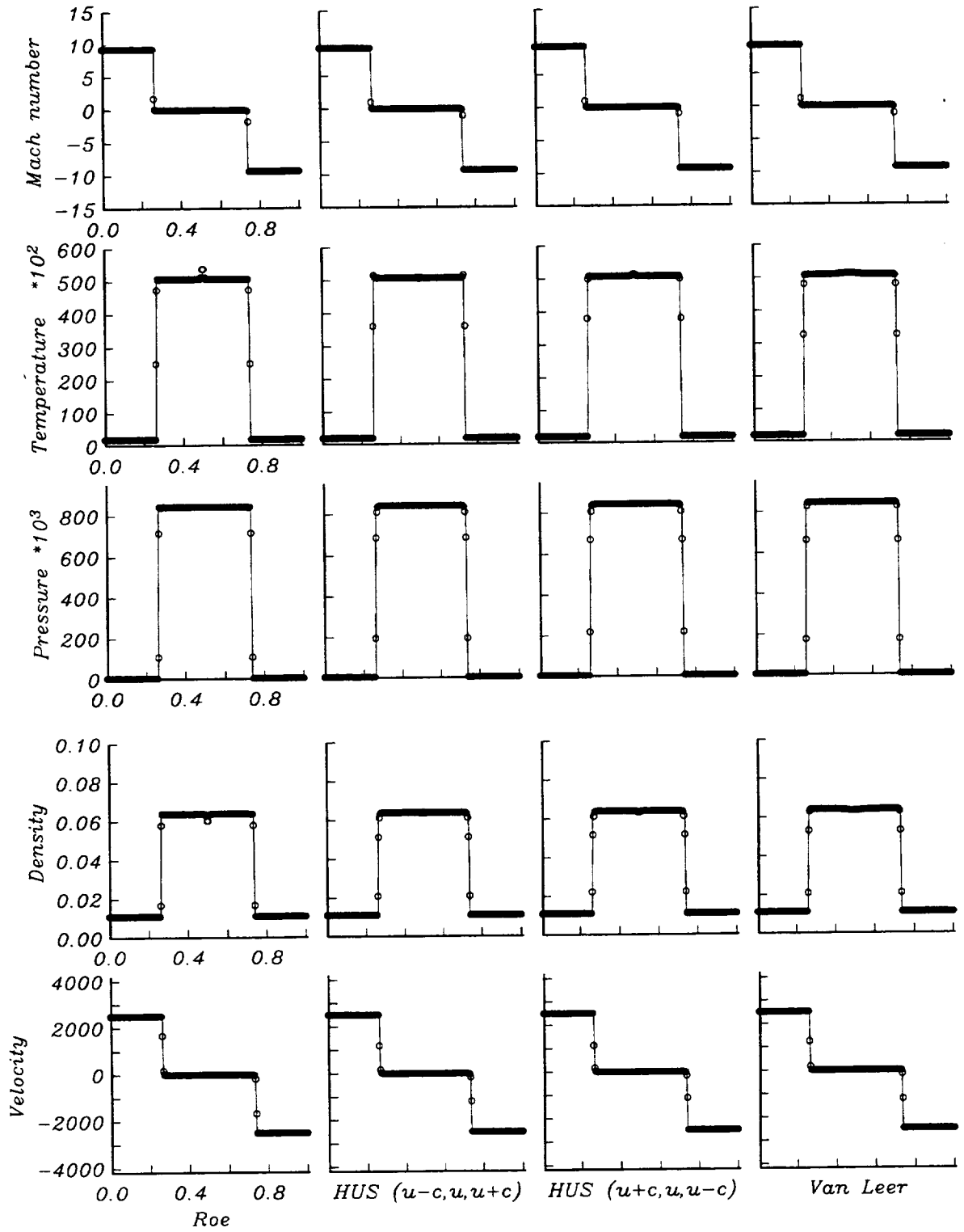


Figure 6. Shock tube problem, Case 5: (a) Comparison of results from the methods of Roe, HUS (Van Leer-Osher hybridization with natural ordering), HUS (Van Leer-Osher hybridization with reverse ordering), Van Leer, and the exact Riemann solver (solid line).

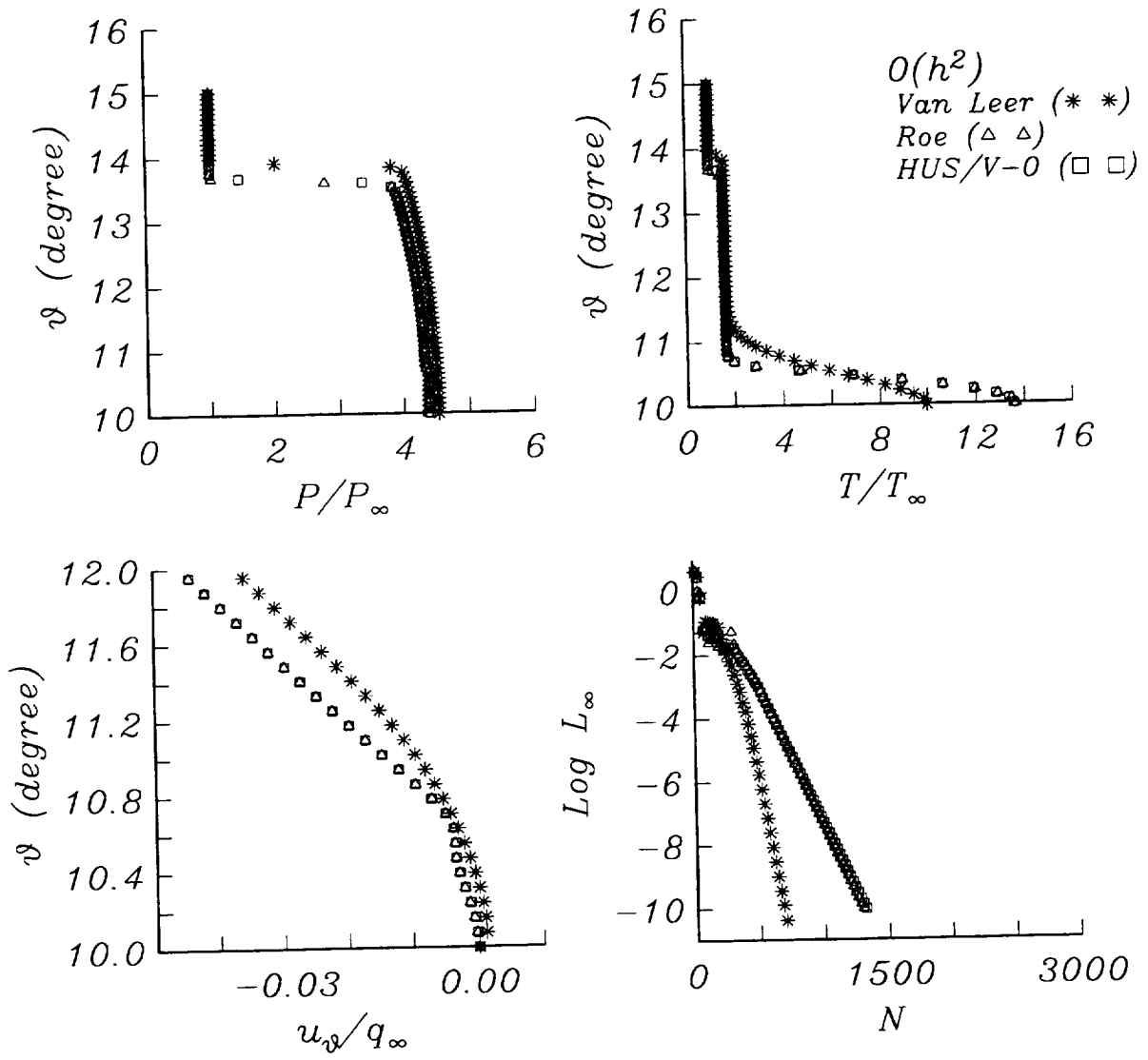


Figure 7. Quasi one-dimensional hypersonic conical flow problem,  $M_\infty = 7.95$ ,  $Re = 4.2 \times 10^5$ , and half cone angle= $10^\circ$ .

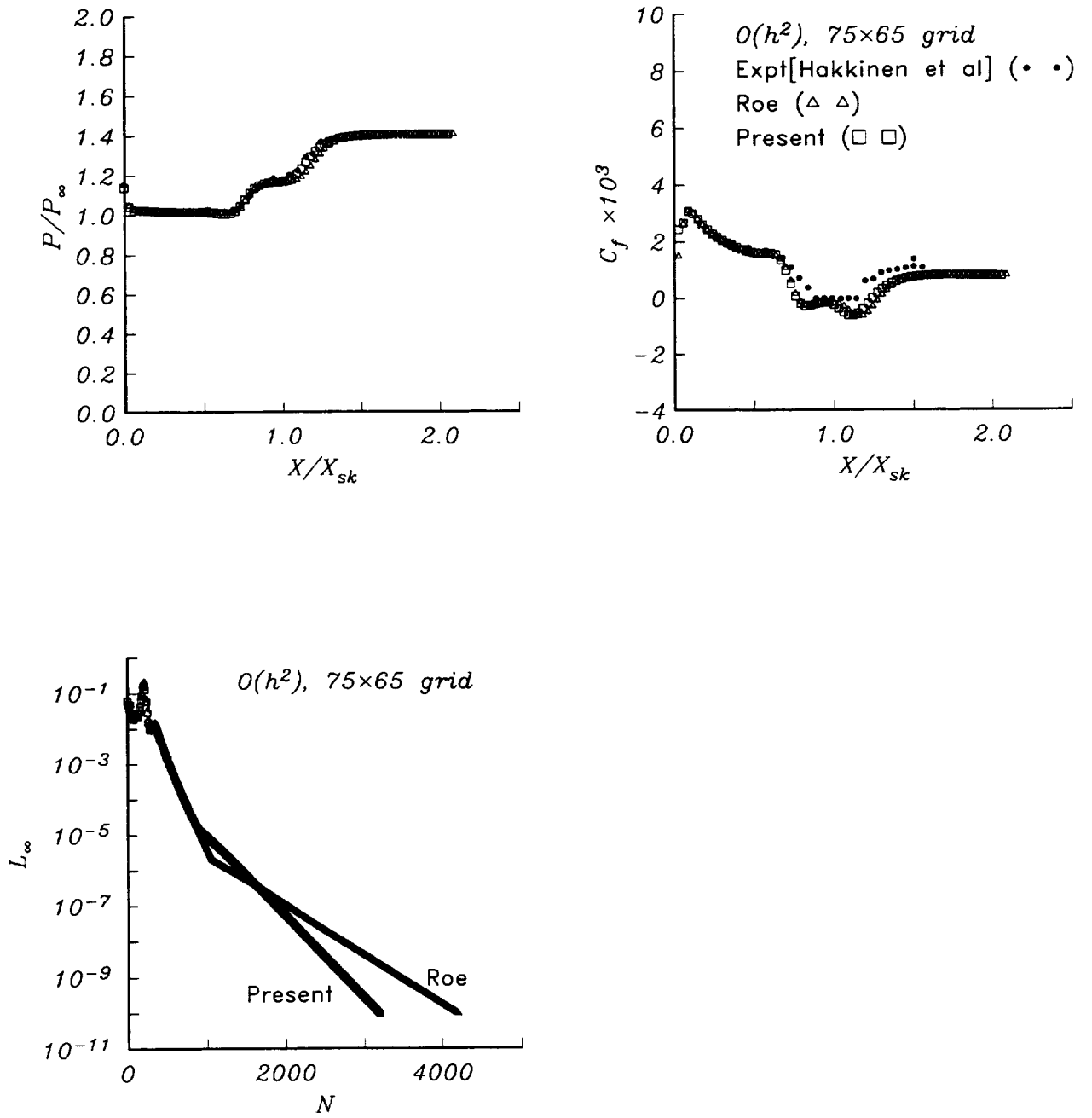


Figure 8. Supersonic shock wave/laminar boundary layer interaction problem: surface pressure, skin-friction coefficient, and convergence history.

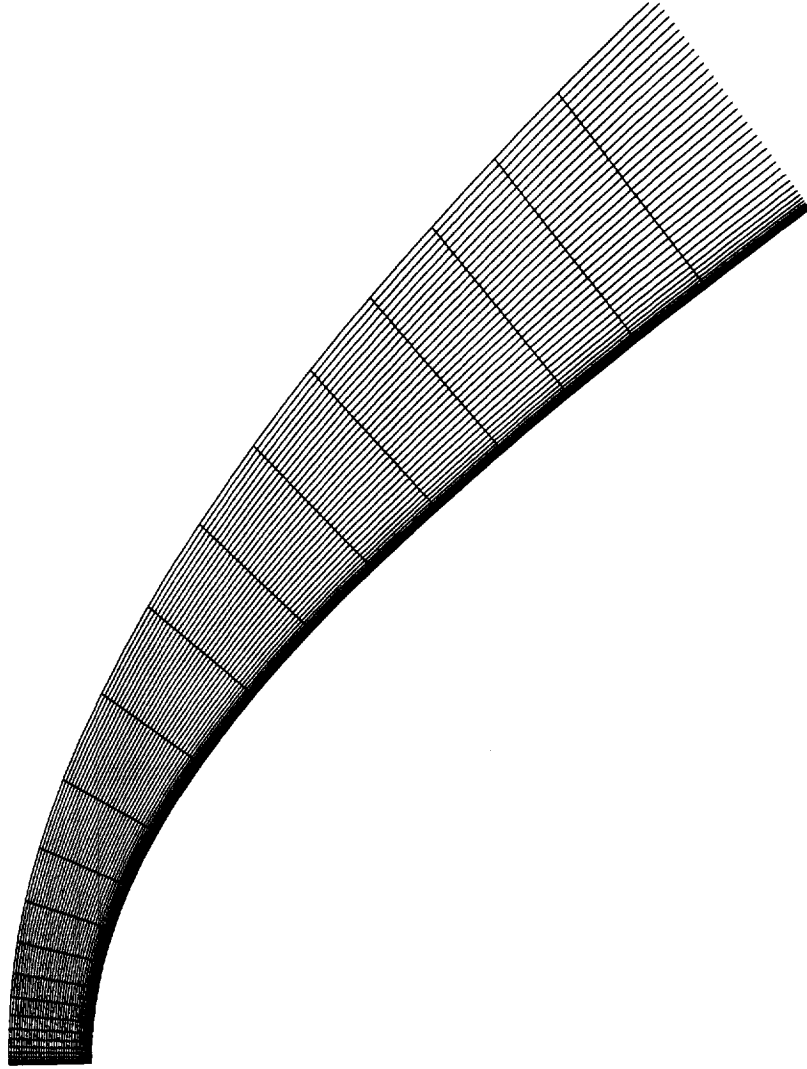


Figure 9. Grid used for hypersonic laminar flow over a hyperboloid :  $M_\infty = 25$ ,  $Re_R = 7684$ ,  $T_\infty = 192.34\text{ K}$ , and  $T_{wall} = 800\text{ K}$ .

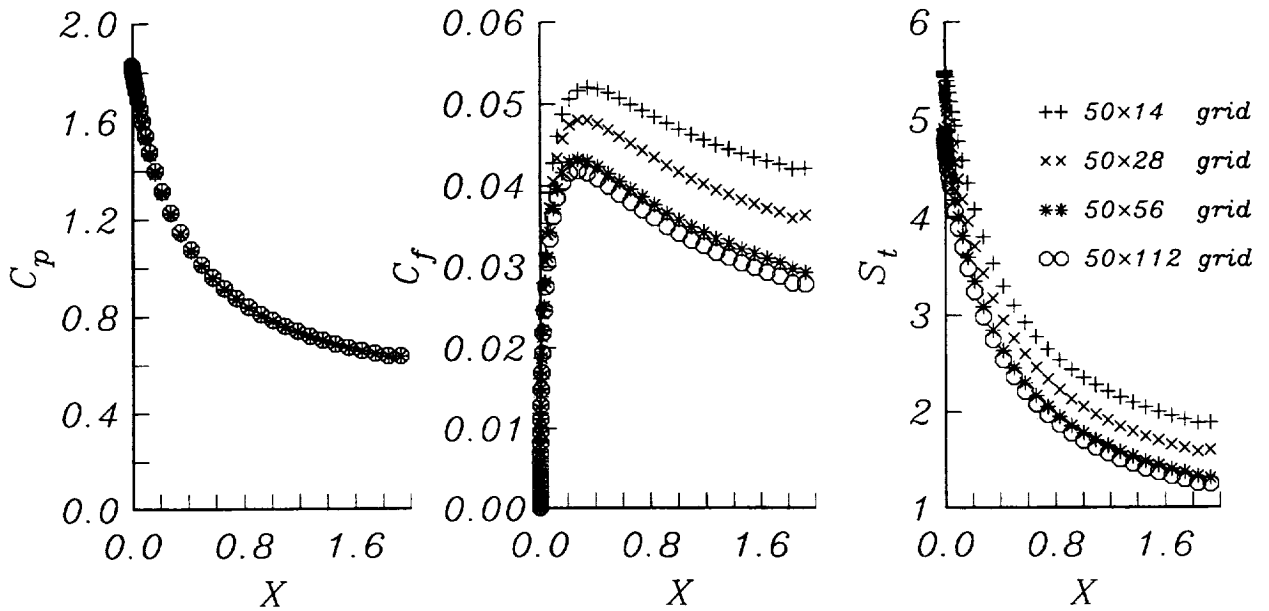


Figure 10. Hypersonic hyperboloid problem: grid refinement study of surface pressure, friction, and heat transfer by the Van Leer method.

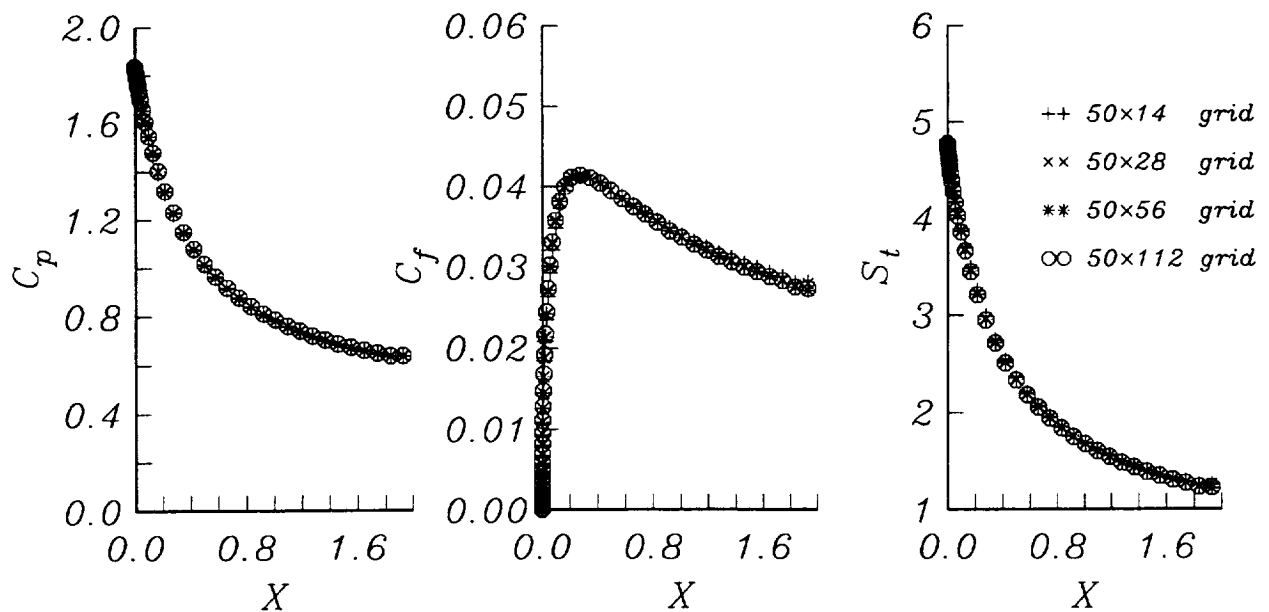


Figure 11. Hypersonic hyperboloid problem: grid refinement study of surface pressure, friction, and heat transfer by the HUS (VL-O) method.

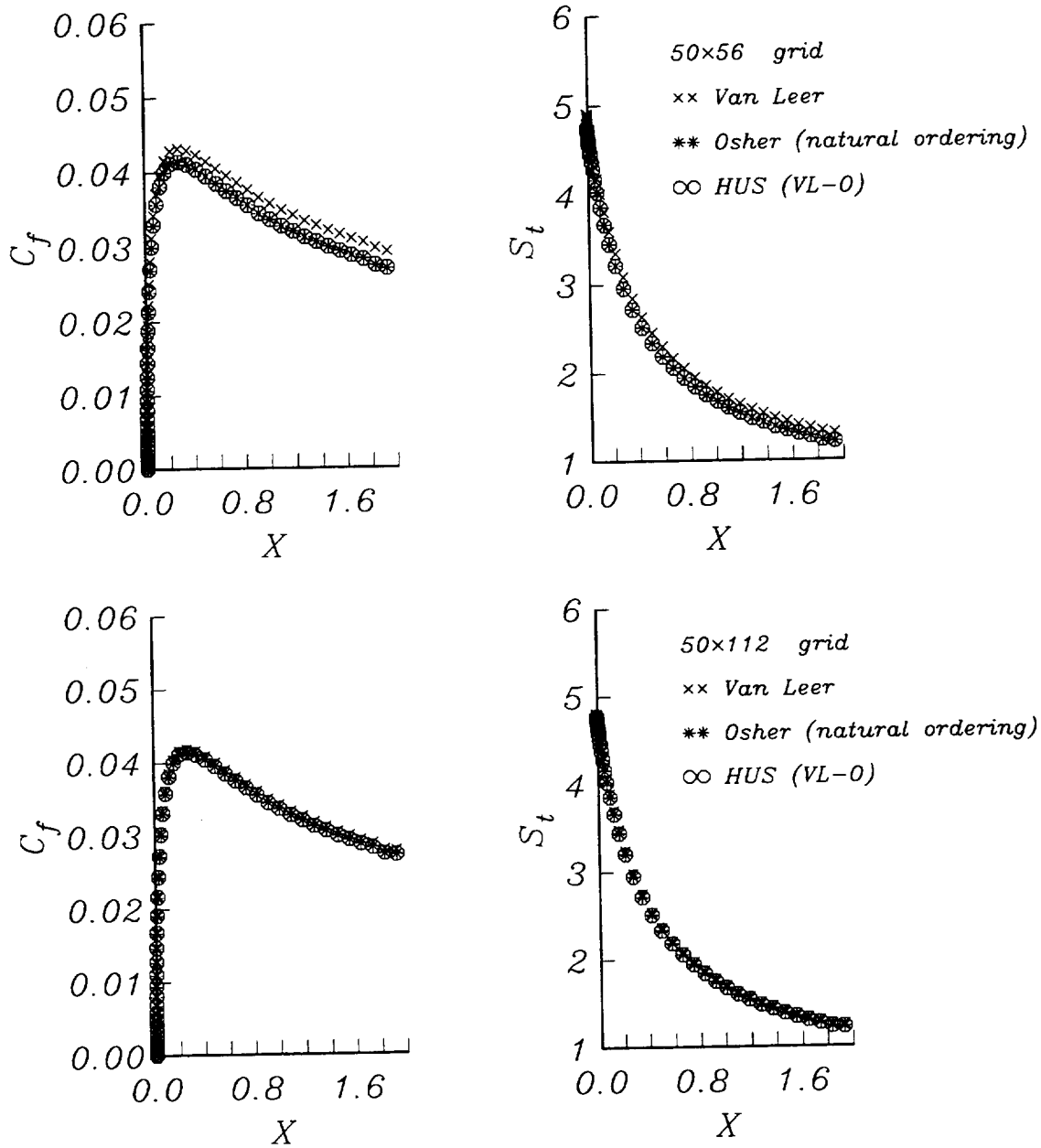


Figure 12. Hypersonic hyperboloid problem: comparison of the Van Leer, Osher, and HUS solutions on surface pressure, friction, and heat transfer on two fine grids.

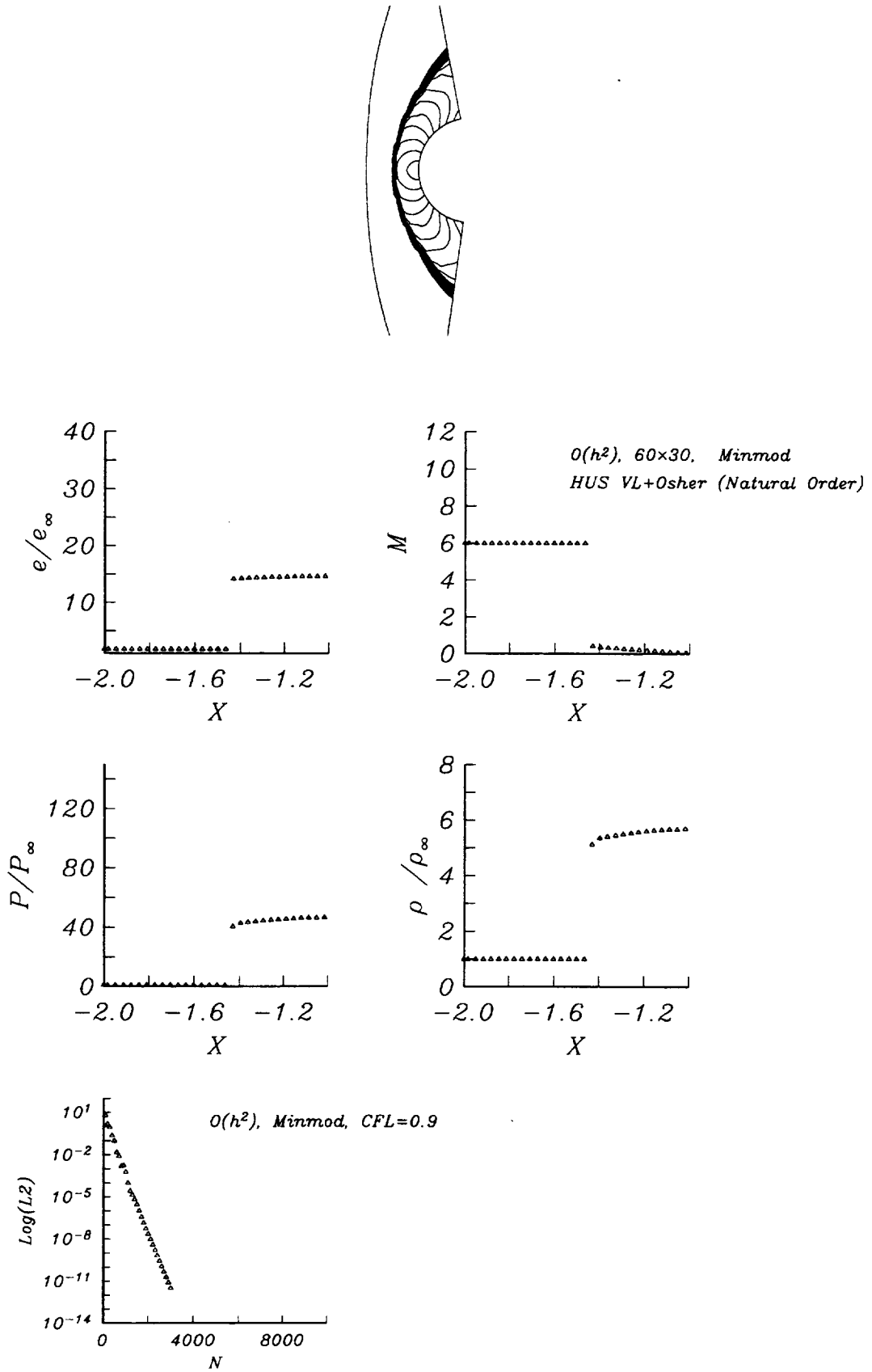


Figure 14. Inviscid HUS solution of *M6* flow over a circular body.



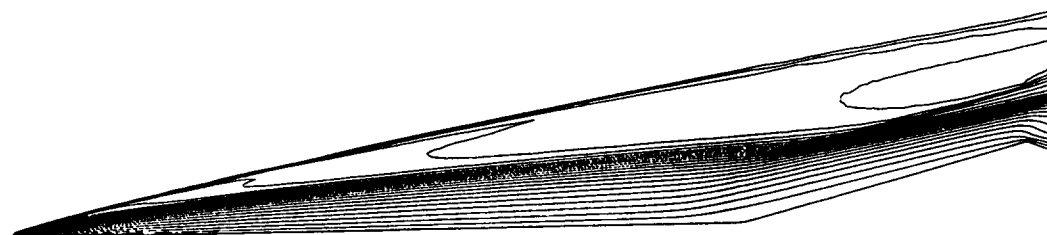
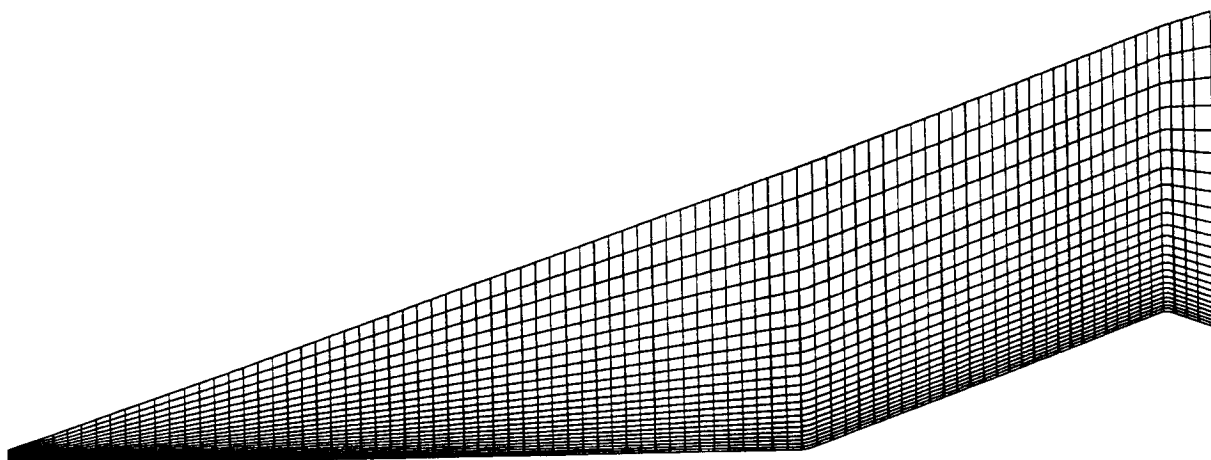


Figure 15. Hypersonic laminar flow over a  $15^\circ$  ramp : (a)  $174 \times 49$  Grid (skipped every other lines for clarity), and (b) Mach number contours .

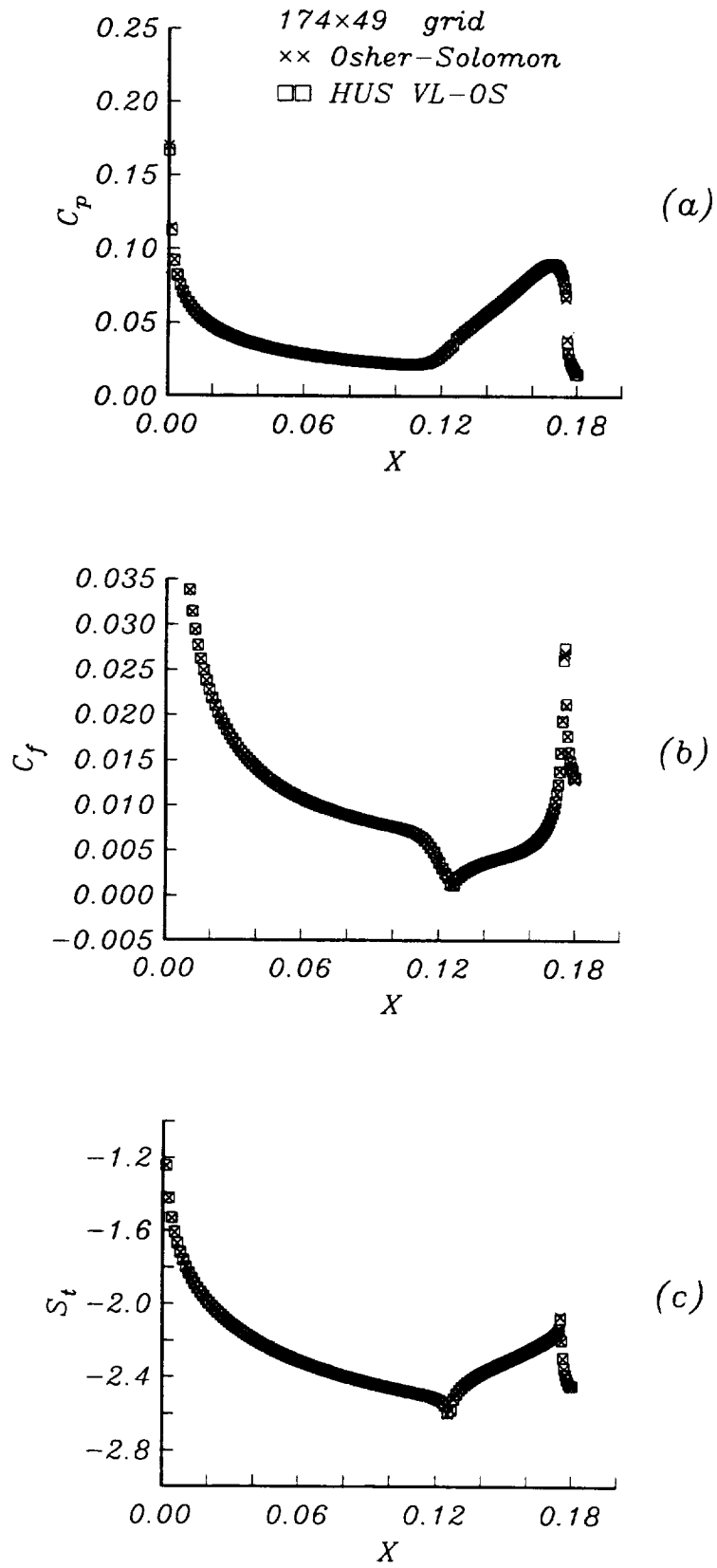


Figure 16. Distribution of surface quantities on the  $15^\circ$  ramp : (a) pressure, (b) friction coefficient, and (c) Stanton number.

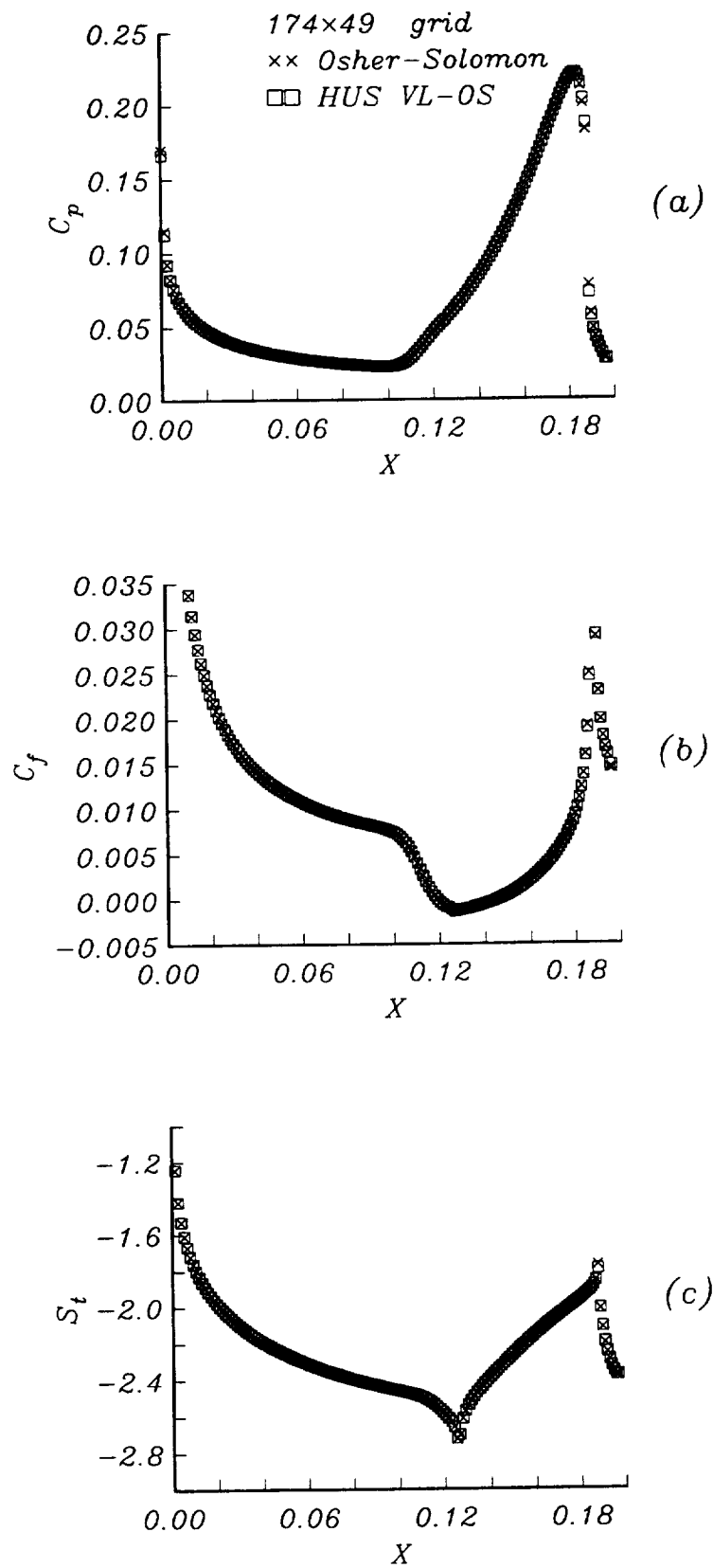


Figure 17. Distribution of surface quantities on the 20° ramp : (a) pressure, (b) friction coefficient, and (c) Stanton number.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE January 1995		3. REPORT TYPE AND DATES COVERED Technical Memorandum
4. TITLE AND SUBTITLE  Hybrid Upwind Splitting (HUS) by a Field-by-Field Decomposition			5. FUNDING NUMBERS  WU-505-62-52	
6. AUTHOR(S)  Frédéric Coquel and Meng-Sing Liou				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135-3191			8. PERFORMING ORGANIZATION REPORT NUMBER  E-9408	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  National Aeronautics and Space Administration Washington, D.C. 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER  NASA TM-106843 ICOMP-95-2	
11. SUPPLEMENTARY NOTES Frédéric Coquel, Institute for Computational Mechanics in Propulsion, NASA Lewis Research Center (work funded under NASA Cooperative Agreement NCC3-233), and Meng-Sing Liou, NASA Lewis Research Center. ICOMP Program Director, Louis A. Povinelli, organization code 2600, (216) 433-5818.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT  Unclassified - Unlimited Subject Categories 02, 34, and 64  This publication is available from the NASA Center for Aerospace Information, (301) 621-0390.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) We introduce and develop in this paper a new approach for upwind biasing: the <i>Hybrid Upwind Splitting</i> (HUS) method. This original procedure is based on a suitable hybridization of current prominent <i>Flux Vector Splitting</i> (FVS) and <i>Flux Difference Splitting</i> (FDS) methods. The HUS method is designed to naturally combine the respective strengths of the above methods while excluding their main deficiencies. Specifically, the HUS strategy yields a family of upwind methods that exhibit the <i>robustness of FVS schemes in the capture of nonlinear waves</i> and the <i>accuracy of some FDS schemes in the resolution of linear waves</i> . We give a detailed construction of the HUS methods following a <i>general and systematic procedure</i> directly performed at the basic level of the <i>Field by Field</i> (i.e. waves) decomposition involved in FDS methods. For such a given decomposition, each field is endowed either with FVS or FDS numerical fluxes, depending on the nonlinear nature of the field under consideration. Such a design principle is made possible thanks to the introduction of a convenient formalism that provides us with a unified framework for upwind methods. The HUS methods we propose bring significant improvements over current methods in terms of accuracy and robustness. They yield <i>entropy-satisfying</i> approximate solutions as they are strongly supported in numerical experiments. Field by field hybrid numerical fluxes also achieve <i>fairly simple and explicit</i> expressions and hence require a computational effort between that of the FVS and FDS. Several numerical experiments ranging from stiff 1D shock-tube to high speed viscous flows problems are displayed, intending to illustrate the benefits of the present approach. We shall assess in particular the relevance of our HUS schemes to viscous flow calculations.				
14. SUBJECT TERMS  Stable, accurate, and efficient hybrid upwind scheme; Navier-Stokes and Euler equations			15. NUMBER OF PAGES 91	
			16. PRICE CODE A05	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	